

From: Mertz, Prema
Sent: Saturday, October 19, 2002 2:53 PM
To: STIC-Biotech/ChemLib
Subject: 09/851,595

Please search SEQ ID NO:11 with protein databases.

Thanks,

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

10/23/02
ABH
1-AA

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52 ; Search time 37 Seconds
(without alignments)
2902.929 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068
Sequence: 1 MPSPGRLALMLCALACASR.....GGLSGGGGFGSGLAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4067	80.2	847	20 AAY53575	Human gonadotropin
2	2886.5	57.0	691	22 AAG09314	Novel human diapo
3	2525.5	49.8	907	20 AAW93889	Human Hg38 protein
4	2525.5	49.8	907	21 AAY90682	Human G protein-co
5	2519.5	49.7	907	21 AAY90687	Human mutant G pro
6	2450	48.3	497	20 AAY53574	Human RBGAP polype
7	2035.5	40.1	951	22 AAB68875	A human guanosine
8	2033	40.1	395	22 AAG67556	Human AOMF05 prote
9	2030.5	40.1	949	20 AAW93904	Human AOMF05 prote
10	2030.5	40.1	949	20 AAW93905	Human AOMF05 prote
11	2030.5	40.1	951	20 AAW93965	Human AOMF05 prote

12	2030.5	40.1	951	20 AAW93906	Human AOMF05 prote
13	2030.5	40.1	951	22 AAY32972	Novel human secret
14	2027.5	40.0	951	20 AAY42168	Human LGR4 protein
15	1996	39.4	692	20 AAW93890	Human HG38 protein
16	1901	37.5	693	20 AAY42169	Human LGR5 protein
17	1367	27.0	266	21 AAB41583	Human ORFX ORF1347
18	1206	23.8	230	22 AAB95066	Human protein sequ
19	1140	22.5	644	20 AAW82318	Human 7-transmembr
20	1102	21.7	631	20 AAY33767	Portion of Tango-7
21	1010.5	19.9	1300	22 ABB64083	Drosophila melanog
22	901	17.8	171	20 AAY53571	Human gonadotropin
23	688.5	13.6	861	22 AAB86961	D. melanogaster pe
24	536	10.6	141	22 ABB11736	Human Leu-rich rep
25	528.5	10.4	699	22 AAR30517	N-terminal of LH r
26	525	10.4	120	22 ABB03064	Human expressed po
27	525	10.4	120	22 AAY20415	Human secreted pro
28	523	10.3	695	13 AAR27558	FSHR. Homo sapien
29	519.5	10.3	692	11 AAR08038	Rat testicular lut
30	518	10.2	695	14 AAW14782	FSH receptor. Hom
31	518	10.2	689	14 AAR30509	N-terminal of LH r
32	517.5	10.2	695	14 AAR30524	N-terminal of LH r
33	517	10.2	695	14 AAR30511	N-terminal of LH r
34	515.5	10.2	695	14 AAR30516	N-terminal of LH r
35	512	10.1	700	14 AAR30516	N-terminal of LH r
36	511	10.1	692	14 AAR30503	N-terminal of LH r
37	510	10.1	692	14 AAR30525	N-terminal of LH r
38	508.5	10.0	695	14 AAR30505	N-terminal of LH r
39	507	10.0	698	14 AAR30523	Human gonadotropin
40	502.5	9.9	202	20 AAY53573	N-terminal of LH r
41	497	9.8	696	14 AAR30523	N-terminal of LH r
42	492.5	9.7	764	12 AAR12504	Canine thyroid scl
43	491.5	9.7	829	22 ABB63439	Drosophila melanog
44	489.5	9.7	696	14 AAR30519	N-terminal of LH r
45	486.5	9.6	696	14 AAR30526	N-terminal of LH r

ALIGNMENTS

RESULT 1
AAV53575
ID AAV53575 standard; Protein: 847 AA.
AC AAY53575;
DT 15-FEB-2000 (first entry)
XX
DE Human gonadotropin receptor partial sequence #5.
KW Human: gonadotropin receptor; screening; ligand; biomedical research;
KW biochemical research; drug; hormone; reproductive tissue; infertility;
KW contraception.
OS Homo sapiens.
XX
XX EP950711-A2.
PN 20-OCT-1999.
PD
XX
XX 02-FEB-1999; 99EP-0200303.
XX
XX 06-FEB-1998; 98EP-0200357.
PR 27-JUL-1998; 98EP-0202519.
PR 24-SEP-1998; 98EP-0203213.
XX
XX (ALXU) AKZO NOBEL NV.
XX
XX Van Der Spek PT, Heikoop JC;
XX WPI; 1999-563673/48.
XX N-FSDB; AAZ40461.
XX New 7 transmembrane gonadotropin receptors, useful for screening for

PT hormone analogs and drugs -

PS Claim 1: Page 24-28; 38pp; English.

XX
XX
CC Sequences AAY53571-Y53578 represent fragments of 3 novel human
CC gonadotropin receptors. The novel gonadotropin receptors can be used to
CC screen for ligands of the receptors. This screen may be used in
CC biomedical and biochemical research to develop new drugs targeted to the
CC gonadotropin receptors. For example, hormone analogs which activate or
CC inhibit the function of the gonadotropin receptors or classical
CC gonadotropin receptors may be detected. The altered expression or
CC dysfunction of the gonadotropin receptors causes conditions in
CC reproductive tissues. The ligands can be used for the treatment of
CC infertility or for contraception.

SQ Sequence 847 AA;

Query Match Best Local Similarity 80.2%; Score 4067; DB 20; Length 847;
Matches 789; Conservative 0; Mismatches 2; Indels 96; Gaps 1;

QY 72 DLSMNNLTLEDPGLFHHIRFLLEELRLSGNHLSHIPGQAFSGSLKILIMONNOGGIPA 131
DB 29 DLSMNNLTLEDPGLFHHIRFLLEELRLSGNHLSHIPGQAFSGSLKILIMONNOGGIPA 88
QY 132 EALMELPSLOSRLDANLISLPERSEGLSLRHLMDLNDALTEIPVALNMLPALQAM 191
DB 89 EALMELPSLOS-----
QY 192 TLALNRSHTPDVAFONLISLVYLHNNRIQHLGTHSEGLHNLFTLNLNKKLOEPPV 251
DB 100 -----
QY 252 AIRTLGRLOELGFHNNNIKAIPKAFMGNPILOTIFYNDPIQFGRSAPFYLPRKLTLS 311
DB 113 AIRTLGRLOELGFHNNNIKAIPKAFMGNPILOTIFYNDPIQFGRSAPFYLPRKLTLS 172
QY 312 LAGANDIOEPDLKGTSTLEITLTPRAGIRLLPSGMCQOLPRLRYLELSHNOIEELPSLH 371
DB 173 LAGANDIOEPDLKGTSTLEITLTPRAGIRLLPSGMCQOLPRLRYLELSHNOIEELPSLH 232
QY 372 RCQKLEELIGLQHNRIWEIGADTFSSLSQALDLSWNAIRSHHPAFSTLSLVKLDLTD 431
DB 233 RCQKLEELIGLQHNRIWEIGADTFSSLSQALDLSWNAIRSHHPAFSTLSLVKLDLTD 292
QY 432 NQUTLPLAGLGLMHLKLNALISOAFSKDSEPKLILEVPAYOCCPYGKMSAFKFA 491
DB 293 NQUTLPLAGLGLMHLKLNALISOAFSKDSEPKLILEVPAYOCCPYGKMSAFKFA 352
QY 492 SGOMEAEDLHLDDESSKRPGLLAROENHYDODDELQLEMEDSKPHPSVOCSPTRGP 551
DB 353 SGOMEAEDLHLDDESSKRPGLLAROENHYDODDELQLEMEDSKPHPSVOCSPTRGP 412
QY 552 FKPEEYLFESGIRLAWAIVLVSLVLCNGLVLLTFVAGGPAFLPYKVVGAIGAGANTLT 611
DB 413 FKPEEYLFESGIRLAWAIVLVSLVLCNGLVLLTFVAGGPAFLPYKVVGAIGAGANTLT 472
QY 612 GISCGLLASVDALTFQGFSEYGARWETGLGCRATNGFLAVIGSASVLLTLAAGVCSVS 671
DB 473 GISCGLLASVDALTFQGFSEYGARWETGLGCRATNGFLAVIGSASVLLTLAAGVCSVS 532
QY 672 SCVRATGKSPISLGSVAGVYGLALAGLAALPLASVGEYASPLCLPAPEGGOPALAG 731
DB 533 SCVRATGKSPISLGSVAGVYGLALAGLAALPLASVGEYASPLCLPAPEGGOPALAG 592
QY 732 FTVAVLVMNSFCFLVAVAGYIKLYCDLPRGDEAVWDCAMVHVAMVLFADGLLYCPVAF 791
DB 593 FTVAVLVMNSFCFLVAVAGYIKLYCDLPRGDEAVWDCAMVHVAMVLFADGLLYCPVAF 652
QY 792 LSEFASMLGLFPYTPRAVSVLLVLPDLPAICNPLILYLFNPFERDRLRLPRAGDSPL 851
DB 653 LSEFASMLGLFPYTPRAVSVLLVLPDLPAICNPLILYLFNPFERDRLRLPRAGDSPL 712

QY 852 AYAAGELEKSSCDSTQALVAFSDVLTLEASSEAGRPGLTETYPSTYTLISCOQPGAPR 911
DB 713 AYAAGELEKSSCDSTQALVAFSDVLTLEASSEAGRPGLTETYPSTYTLISCOQPGAPR 772
QY 912 LEGSHCVPEENHNGNPPSPMDGELLRLRBSSTRPAGGSLGGGFGFP 958
DB 773 LEGSHCVPEENHNGNPPSPMDGELLRLRBSSTRPAGGSLGGGFGFP 819

RESULT 2

ID ABG09314 standard; Protein; 691 AA.

AC ABG09314;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9305.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

OS food supplement; medical imaging; diagnostic; genetic disorder.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PA 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

XX N-PSDB; AAS73501.

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PS Claim 20; SEQ ID No 39673; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations in

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. ABG00010-ABG30377 represent novel human

XX diagnostic amino acid sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at [ftp.wipo.int/publ/published_sequences](http://wipo.int/publ/published_sequences).

SQ Sequence 691 AA;

Query Match Best Local Similarity 57.0%; Score 2886.5; DB 22; Length 691;
Matches 789; Conservative 0; Mismatches 2; Indels 96; Gaps 1;


```

Db 728 ALILNLSLCEFLMTIATYKLYCNLDKGLDENIMDCSWVKHIALLEFLNCLNCPVAFILSF 787
OY 795 ASWLGLEPVPPEAVKSVLIVLPLPACLNPLLYLLNPHFRDRLRKPRA-----GDSG 849
Db 788 SSLINLTFISPEYIKFILLVWVPLPACLNPLLYLLNPHFRDRLRKPRA-----GDSG 849
OY 850 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEAGRP-----PGLETYGFPSVTLLSC 904
Db 848 SLMSINSDVEKQSCDSTQALVAFSDVDLILEASEAGRP-----PGLETYGFPSVTLLSC 904

```

RESULT 4

AAI90682

ID AAY90682 standard; Protein; 907 AA.

AC AAY90682;

DT 21-AUG-2000 (first entry)

DE Human G protein-coupled receptor HG38.

KW G protein-coupled receptor; GPCR; constitutively active;

KW intracellular loop 3; transmembrane domain 6; drug screening;

KW agonist; antagonist.

OS Homo sapiens.

PN MO200022129-A1.

PD 20-APR-2000.

PF 12-OCT-1999; 99WO-US23938.

PR 13-OCT-1998; 98US-0170496.

PA (AREN-) ARENA PHARM INC.

PI Behan DP, Chalmers DT, Liaw CW;

DR WPI; 2000-329165/28.

DR N-PSDB; AAA30770.

Non-endogenous constitutively activated human G protein-coupled
receptors, useful for identifying agonists for use as pharmaceutical
agents

Example 1; Page 317-320; 341pp; English.

The invention relates to constitutively active, non-endogenous versions
of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
AAY90677 and AAY90683-190687), and to DNA encoding them (AAA30709-A30743
and AAA30775-A30779). The mutant proteins of the invention contain a
mutation in a portion of the protein comprising intracellular loop 3
(IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
is substituted for an endogenous residue in IC3 at a position 16 amino
acids N-terminal of an endogenous proline in TM6 to form a sequence
X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg
or Ala, and is preferably Lys. When the endogenous residue at this
position is Lys, this residue is replaced by His, Arg or preferably Ala.
The 15 amino acid stretch between the substituted amino acid and the Pro
may be endogenous, non-endogenous, or a mixture of endogenous and
non-endogenous residues. The constitutively active GPCRs are useful for
identifying antagonists, agonists and partial agonists for use as
pharmaceutical agents. The mutant proteins are also useful in research
settings for elucidating the roles of the receptors in normal and
diseased conditions. Antagonists for a particular GPCR are useful for
treating diseases and disorders associated with that receptor. Because
the novel mutant GPCRs are constitutively active, they can be used
directly for screening of compounds without the need for endogenous
ligands. The present sequence represents a human wild-type GPCR referred
to in an exemplification of the invention.

SQ Sequence 907 AA;

Query Match 49.8%; Score 2525.5; DB 21; Length 907;
Best Local Similarity 55.3%; Pred. No. 6e-197;
Matches 497; Conservative 136; Mismatches 235; Indels 31; Gaps 9;

```

OY 23 GGAPOGPETACPARCOEDG-IMLSADCSLGLSAMPBGDIPPLVATLIDLSMNLTEL 81
Db 22 GSSPRGCVLLRGCPTRCHCEPEDEPDRMLRVDCSDGLSELPSNLVSFTSYLDLSMNNISOL 81
OY 82 QPGLFHLRFELEELRSGNHLSHIIPGAFSGYSLKIMONNOLOGIPAEALMEPLSQ 141
Db 82 LRPPLPSLFLLEELRLAGNALITYPKGAFGLSLKLVLMONNOHLHVPEALQNRSLQ 141
OY 142 SLRLDANLISLVERSEFEGSLSLRLMLDNLALTEIPVARNLNPALQAMTLALNRSHI 201
Db 142 SLRLDANLISLVERSEFEGSLSLRLMLDNLALTEIPVARNLNPALQAMTLALNRSHI 201
OY 202 PDYAFONLTSIVVHLHNNRIQHLGTHSEFGLNLETLDLNTNKKLOEPFVAIRTGLOE 261
Db 202 PDYAFONLTSIVVHLHNNRIQHLGTHSEFGLNLETLDLNTNKKLOEPFVAIRTGLOE 261
OY 262 LGFHNNNIKAIFPKAFMGNPLQTIHFYDNPLOFVGRSAFOYLPRKHTLSLNGAMDIQOE 321
Db 262 LGFHNNNIKAIFPKAFMGNPLQTIHFYDNPLOFVGRSAFOYLPRKHTLSLNGAMDIQOE 321
OY 322 PDLKGTSLLEITLRAGIRLLPSGMCQOLPRLRVLELSHNOIEELPSLHRCQKLEELGL 381
Db 322 PDLKGTSLLEITLRAGIRLLPSGMCQOLPRLRVLELSHNOIEELPSLHRCQKLEELGL 381
OY 382 QHNRIMEIGADPFSSQSLSDALDLSMNAIRSHIPAFSTHSLVLTLDLNDQTLPLAG 441
Db 382 RHNELYEIKVDPFQOLLSLRSLNLMNKIATIHHPAFSTHSLVLTLDLNDQTLPLAG 441
OY 442 LGLMLHLKGNLALISQAFSKDSFKRLILEVRYAYOCCPYGMCASFKASGOW----- 495
Db 442 LGLMLHLKGNLALISQAFSKDSFKRLILEVRYAYOCCPYGMCASFKASGOW----- 495
OY 496 EAEDLHDESSSKRPLDLLARQANHNHDQDLDELQLEM-EDSKHSPVOCSPYPPGPKP 554
Db 496 EAEDLHDESSSKRPLDLLARQANHNHDQDLDELQLEM-EDSKHSPVOCSPYPPGPKP 554
OY 502 SMDLH-----KKDAGMFOAQDE---RLEDELDFEEDLKLHVSQSPSPGPKP 550
Db 502 SMDLH-----KKDAGMFOAQDE---RLEDELDFEEDLKLHVSQSPSPGPKP 550
OY 555 CEYLFEESWIGRLAWAIVLISVLCNGIVLIFYFAGGPARLPVYKRVGAIAGANTLTGIS 614
Db 555 CEYLFEESWIGRLAWAIVLISVLCNGIVLIFYFAGGPARLPVYKRVGAIAGANTLTGIS 614
OY 551 CEHLIDGWLIRIGWTLVAVLALTCNALVTSTYFR-SPLIYSPKILLIVIAVNMILGVS 609
Db 551 CEHLIDGWLIRIGWTLVAVLALTCNALVTSTYFR-SPLIYSPKILLIVIAVNMILGVS 609
OY 615 CGLASVDALFFGOFSEYEGARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSVSCV 674
Db 615 CGLASVDALFFGOFSEYEGARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSVSCV 674
OY 610 SAVLAGVDAPFFGSPARIGAMWENGCHVIGFLSIFESSEVFLTLAALREGFVYKS 669
Db 610 SAVLAGVDAPFFGSPARIGAMWENGCHVIGFLSIFESSEVFLTLAALREGFVYKS 669
OY 675 RAYKSPSLGSVRAGVIGCLALAGLAALPLASVGEASPLCIPYAPPEQPAALGFTV 734
Db 675 RAYKSPSLGSVRAGVIGCLALAGLAALPLASVGEASPLCIPYAPPEQPAALGFTV 734
OY 670 AKFETKAPFSSLKAYIILLCALALITMAAVPLGSKTGASPLCLP--PREPSTMGIMV 727
Db 670 AKFETKAPFSSLKAYIILLCALALITMAAVPLGSKTGASPLCLP--PREPSTMGIMV 727
OY 735 ALVMAISFCLPVAVAGVITKLYCDLPRGDPEAVWDCAMVRHVAWITFADGLICVAVLSF 794
Db 735 ALVMAISFCLPVAVAGVITKLYCDLPRGDPEAVWDCAMVRHVAWITFADGLICVAVLSF 794
OY 728 ALILNLSLCEFLMTIATYKLYCNLDKGLDENIMDCSWVKHIALLEFLNCLNCPVAFILSF 787
Db 728 ALILNLSLCEFLMTIATYKLYCNLDKGLDENIMDCSWVKHIALLEFLNCLNCPVAFILSF 787
OY 795 ASWLGLEPVPPEAVKSVLIVLPLPACLNPLLYLLNPHFRDRLRKPRA-----GDSG 849
Db 795 ASWLGLEPVPPEAVKSVLIVLPLPACLNPLLYLLNPHFRDRLRKPRA-----GDSG 849
OY 850 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEAGRP-----PGLETYGFPSVTLLSC 904
Db 850 PLAYAAAGELKSSCDSTQALVAFSDVDLILEASEAGRP-----PGLETYGFPSVTLLSC 904

```

RESULT 5

AAI90687

ID AAY90687 standard; Protein; 907 AA.

AC AAY90687;

DT 21-AUG-2000 (first entry)

XX Human mutant G protein-coupled receptor HG38 (V765X).

DE G protein-coupled receptor: GPCR; constitutively active;

XX intracellular loop 3; transmembrane domain 6; drug screening;

KW agonist; antagonist; mutant; mutein.

XX Homo sapiens.

OS Synthetic.

XX W0200022129-A1.

PN 20-APR-2000.

PD 12-OCT-1999; 99WO-US23938.

PF 13-OCT-1998; 98US-0170496.

PR (AREN-) ARENA PHARM INC.

XX Behan DP, Chalmers DT, Liaw CW;

PI WPI: 2000-329165/28.

DR N-PSDB; AAA30779.

XX Non-endogenous constitutively activated human G protein-coupled

PT receptors, useful for identifying agonists for use as pharmaceutical

PT agents -

XX Example 2; Page 332-335; 341pp; English.

PS The invention relates to constitutively active, non-endogenous versions

XX of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-

CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743

CC and AAA30775-A30779). The mutant proteins of the invention contain a

CC mutation in a portion of the protein comprising intracellular loop 3

CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,

CC is substituted for an endogenous residue in IC3 at a position 16 amino

CC acids N-terminal of an endogenous proline in TM6 to form a sequence

CC X-(AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg

CC or Ala, and is preferably Lys. When the endogenous residue at this

CC position is Lys, this residue is replaced by His, Arg or preferably Ala.

CC The 15 amino acid stretch between the substituted amino acid and the Pro

CC may be endogenous, non-endogenous, or a mixture of endogenous and

CC non-endogenous residues. The constitutively active GPCRs are useful for

CC identifying antagonists, agonists and partial agonists for use as

CC pharmaceutical agents. The mutant proteins are also useful in research

CC settings for elucidating the roles of the receptors in normal and

CC diseased conditions. Antagonists for a particular GPCR are useful for

CC treating diseases and disorders associated with that receptor. Because

CC the novel mutant GPCRs are constitutively active, they can be used

CC directly for screening of compounds without the need for endogenous

CC ligands. Sequences AAY90643- AAY90677 and AAY90683-Y90687 the mutant

CC human GPCRs of the invention.

XX Sequence 907 AA:

SD

Query Match 49.7%; Score 2519.5; DB 21; Length 907;

Best Local Similarity 55.2%; Pred. No. 1.9e-196;

Matches 496; Conservative 136; Mismatches 236; Indels 31; Gaps 9;

QY 23 GGAQPGPGTACPAPOCHQEDG-IMLSADCSSELSAVGDDPTLTAADLSMNNITEL 81

DB 22 GSSPSRSVLLRGCPTRHCEPDGRMLRLVDCSDGLSELPNSLVFTSYDLSMNNISQL 81

QY 82 QPGLEHRLRLELELRISGNHLSHPGQAFSGLYSLKTIIMLONNOLGGIPAEALWELPSIQ 141

DB 82 LPNLPPLRFLERLRAGNMLTYIPKGAFTGIVSLKVIIMLONNOLRHVPEALONLRISIQ 141

QY 142 SLRDANLISLVPRSEFGLSLRLHMLLDNALTEIPVRLNNIPALQAMTLNLRISHI 201

DB 142 SLRDANHISYVPPSCFGLSHSLRHLMLLDNALTEIPVQAFRSLSLAQAMTLNLRISHI 201

QY 202 PDYAFQNLISLVYLHLHNNRIQHLGHSFEGJLHNETLTDLNANKLOEPPVAIRTLGRIOE 261

DB 202 PDYAFQNLISLVYLHLHNNRIHSLGKCFDGLHSETLTDLNANNDEPFTALRTLSNKE 261

QY 262 LGFHNHNKAIKPKAFMGNDLQTFHYFNPQIOFVGRSAFQYLPRKHTLSLNGADIOFE 321

DB 262 LGFHNHNKIRISIPKAFVGNPSLTIHFYDNPQIOFVGRSAFOHLPELRTLTLNGASQITEF 321

QY 322 PDLKGTTSLEILTLTRAGIRFLPSGMCQQLPRFVLELSHNOIEELPSLRCKOKEIGL 381

DB 322 PDLTGTANLESITLTLGTAQISLSLPQVPCNDLPNLQVLDLSTNLEDDPSVYCCOKLOKIDL 381

QY 382 OHNRIMEIGADTFQSLSQALDLSMMAIRSHIPEAFSTLSLVKIDLTDNOLTLPLAG 441

DB 382 RHNETIETIKVDTFQOQLSLSLSLANLANKIALIHPNMFSTLPSLIKIDLSNLSPTIG 441

QY 442 LGGLMHLKLGKGNLALSGAFSKDSFPKRLILEVYAYQCCPYGKACAFRAGSQ----- 495

DB 442 LHGLTHLKLGKGNHLOSLISSENPPELKIEMPYAVQCAFVCENAYKISNOMKGDNS 501

QY 496 EAEDLHLDDESSKRPGLGLARQAEHNYDQDLDELQLEM-EDSKPHPSVOCSPTPGPFPR 554

DB 502 SMDLH-----KKDAGMEQAODE---RDLEDFLIDPEEDLKALHSVOCSPSPGPFPR 550

QY 555 CEYLFESMGIRLAWAVTLVLSVLCNGILVLTVEAGGPAPLPYKFFVGAAGATLTLGIS 614

DB 551 CEHLDDMLIRIGVITLAVLALCNALVTSVFR-SPLYISPIKLTLGVIAAVNMLTGVS 609

QY 615 CGLIASVDALTFQGFSEYGARWETGLCCRATGFLAVIGSASVILLTLAVOCSSVSCV 674

DB 610 SAVLAGVDFAFTFGSFARHGWENGVCVHGFISIAFSSEVFLTLALERFYSKYIS 669

QY 675 RAYKSPSLGVRAGVGLCGLAALPLASVGEYASPLCLPYAPRPGOPALGFTV 734

DB 676 AKFTKAPFSSLKVTILLCALIALTMAAVPLGSGSKIGASPLCLPL-PPGEPSITMGYV 727

QY 735 ALVMMNSFCPLVAVGAYIKLYCDLPRGDFEAVWDCAMVRHVAWLIFPDGLICVAPLSTF 794

DB 728 ALIILNSICFLMMLTAVTKLYCNDKGLDENIMDCSMKKHIALILFNCLINCVAFLSTF 787

QY 795 ASMGFLPVYDEAVKSVLLVPLPACINPLLYLLEPMPHEDDLRRIRPRA-----GDSG 849

DB 788 SSLIMWFISEVYKIFILLVPLPACINPLLYLLEPMPHEDDLVSLRKQTYVTRSKHP 847

QY 850 PLAYAAAGELEKSSQDSQALVAFSDVDLIEASEAGR-----PGLFTYGFPSVTLISC 904

DB 848 SLMSINSDDVEKQSDSTQALVFTTSSITYDLPPSSVPSPAYVTESCHLSVAFAVPC 906

RESULT 6

AAV53574

ID AAY53574 standard; Protein: 497 AA.

XX AAY53574;

AC

XX

DT 15-FEB-2000 (first entry)

XX

DE Human gonadotropin receptor partial sequence #4.

XX

KW Human; gonadotropin receptor; screening; ligand: biomedical research;

KW biochemical research; drug; hormone; reproductive tissue; infertility;

XX

OS Homo sapiens.

XX

PN EP950711-A2.

PD 20-OCT-1999.

XX

PF 02-FEB-1999; 99EP-0200303.

XX

PR 06-FEB-1998; 98EP-0200357.

PR 27-JUL-1998; 98BP-0202519.
 PR 24-SEP-1998; 98BP-0203213.
 XX
 PA (AKU) AKZO NOBEL NV.
 XX
 PI Van Der Spek PJ, Heikoop JC;
 XX WPI; 1999-563673/48.
 DR N-PSDB; AA40460.
 XX
 PT New 7 transmembrane gonadotropin receptors, useful for screening for
 PT hormone analogs and drugs -
 XX
 PS Claim 1; page 18-21; 38pp; English.
 XX
 CC Sequences AAY53571-Y53578 represent fragments of 3 novel human
 CC gonadotropin receptors. The novel gonadotropin receptors can be used to
 CC screen for ligands of the receptors. This screen may be used in
 CC biomedical and biochemical research to develop new drugs targeted to the
 CC gonadotropin receptors. For example, hormone analogs which activate or
 CC inhibit the function of the gonadotropin receptors or classical
 CC gonadotropin receptors may be detected. The altered expression or
 CC dysfunction of the gonadotropin receptors causes conditions in
 CC reproductive tissues. The ligands can be used for the treatment of
 CC infertility or for contraception.
 CC
 XX
 SQ Sequence 497 AA;
 Query Match 48.3%; Score 2450; DB 20; Length 497;
 Best Local Similarity 99.6%; Pred. No. 3.5e-191;
 Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 490 KASGWEAEDLHLDEESSKRLGLLAROENHYDQDDELQLEMEDSKPHPSVCCSPTP 549
 Db 1 KASGWEAEDLHLDEESSKRLGLLAROENHYDQDDELQLEMEDSKPHPSVCCSPTP 60
 QY 550 GPFKCEYLFESWGRILVMAIVLLSVLCNGVLTFRAGGAPLPYKFFVGAITAGANT 609
 Db 61 GPFKCEYLFESWGRILVMAIVLLSVLCNGVLTFRAGGAPLPYKFFVGAITAGANT 120
 QY 610 LTGISCGLLASVDALTFQGFSEYGARWETGLCCRATGFLAVGSPASVLLTLTAVQCSV 669
 Db 121 LTGISCGLLASVDALTFQGFSEYGARWETGLCCRATGFLAVGSPASVLLTLTAVQCSV 180
 QY 670 SVSCVRATGKSPSLGSAVAGVLCIALAAGLAALPLASVGEYASPLCLPYAPREGOPAA 729
 Db 181 SVSCVRATGKSPSLGSAVAGVLCIALAAGLAALPLASVGEYASPLCLPYAPREGOPAA 240
 QY 730 LGFVALVMMNSFCFLVYVAGVITKLYCDLPGRDPEAVWDCAMVRHVAALIRADGLTYPV 789
 Db 241 LGFVALVMMNSFCFLVYVAGVITKLYCDLPGRDPEAVWDCAMVRHVAALIRADGLTYPV 300
 QY 790 AFLSPASMLGLFVTPPEAVKSVLLVPLPACLNPLLYLTFNPHFRDRLRLRPAGDSG 849
 Db 301 AFLSPASMLGLFVTPPEAVKSVLLVPLPACLNPLLYLTFNPHFRDRLRLRPAGDSG 360
 QY 850 PLAYAAAGELKSSCDSTQALVAFSDVDLIEASEAGRPCELETFYGPSTVLLISCGQPGA 909
 Db 361 PLAYAAAGELKSSCDSTQALVAFSDVDLIEASEAGRPCELETFYGPSTVLLISCGQPGA 420
 QY 910 PRLGSHVCEPEGNHFNPNQPSMDGELLRLRAEGSTPAGGGISGGGGOP 958
 Db 421 PRLGSHVCEPEGNHFNPNQPSMDGELLRLRAEGSTPAGGGISGGGGOP 469
 RESULT 7
 AAB68875
 ID AAB68875 standard; Protein; 951 AA.
 XX
 AC AAB68875;
 XX
 DT 24-APR-2001 (first entry)
 XX

DE Human RECAP polypeptide, SEQ ID NO: 5.
 XX
 XX Human; RECAP; receptors and associated proteins; cerebroprotective;
 KW neurotropic; neuroprotective; anticonvulsant; antiparinsonian; anti-HIV;
 KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
 KW antithyroid; immunosuppressive; nephrotoxic; antiqut; thymimetic;
 KW cytosolic; antibacterial; virucide; fungicide; protozoacide;
 KW antarteriosclerotic; hepatotropic; gene therapy; infection; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200107612-A2.
 XX
 PD 01-FEB-2001.
 XX
 PF 21-JUL-2000; 2000WO-US20035.
 XX
 PR 21-JUL-1999; 99US-0145232.
 PR 07-OCT-1999; 99US-0158578.
 PR 12-NOV-1999; 99US-0165192.
 XX
 PA (INCY) INCYTE GENOMICS INC.
 XX
 PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
 PI Baughn MR, Lu DM, Hillman JL, Patterson C, Lal P;
 DR WPI; 2001-168554/17.
 XX
 DR N-PSDB; AAF58559.
 XX
 PT Novel receptors and associated proteins for diagnosis and treatment of
 PT neurological disorders, immunological disorders including autoimmune/
 PT inflammatory disorders and cell proliferative disorders such as cancer
 PT
 PS Claim 1; page 97-99; 128pp; English.
 XX
 CC The present sequence is a human RECAP (receptors and associated
 CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful
 CC in the diagnosis, treatment and prevention of neurological disorders
 CC such as stroke, Alzheimer's disease, Pick's disease, Huntington's
 CC disease, dementia, Parkinson's disease, Down's syndrome, amyotrophic
 CC lateral sclerosis, multiple sclerosis, bacterial and viral meningitis,
 CC CJD (Creutzfeldt-Jakob disease), GSS (Gerstmann-Strausler-Scheinker
 CC syndrome), immunological disorders, including autoimmune/inflammatory
 CC disorders such as AIDS, digestive's syndrome, severe combined
 CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
 CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
 CC diabetes mellitus, Good pasture's syndrome, gout, Grave's diseases,
 CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
 CC bacterial, fungal, parasitic, protozoal, and helminthic infections, and
 CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis and cancer.
 CC
 XX
 SQ Sequence 951 AA;
 Query Match 40.2%; Score 2035.5; DB 22; Length 951;
 Best Local Similarity 46.5%; Pred. No. 6.4e-157;
 Matches 434; Conservative 129; Mismatches 302; Indels 69; Gaps 12;
 QY 1 MSPPGRLALMLCAALASRRAGAPQPGPTACAPACHQEDGIMLSADCSGLISAV 60
 Db 1 MPGLGLT---LC--FLALGLGSGAGSPCAAPPCSCDGD---RVYDCSGKLAV 51
 QY 61 PGDDPLTAVYLDLMMNLTQLPGLFHLRLTELRLSGNLISHPQASGLYSKILM 120
 Db 52 PEGLSAFTQALDISMNNITQPEDAFKPNPYLEELQAGNLSFIRKALSGELKELVLT 111
 QY 121 LQNNQIGTIPAEALMELPQSLQSLRLDANLISLVERSEFGTSLRHMLDNDNATEIPVR 180
 Db 112 LQNNQIKTVPSEAIRGLSALQSLRLDANLITSVEDEFGELVQRLHMLDNDNATEIPVR 171
 QY 181 ALNNLPALQAMTLALNLSHPYAFQNLISLVYLAHNNRIGHLGTSHSEGLHNEETLD 240
 Db 112 LQNNQIKTVPSEAIRGLSALQSLRLDANLITSVEDEFGELVQRLHMLDNDNATEIPVR 171


```

Db 172 PLSNIPQLALTLANKISSIPDEAFNLISLVYLHLNNKTRLSQHCDFGLDNLNLTID 231
QY 241 LYNKRLQEPVAIRTLQLOELGFHNNIKALPEKAFMGNPLLOTIHEYDNDIQEVRSA 300
Db 232 LNNYNLGEFPAQIKALPSLKELGFSNSISVYPCGAFDQNPILLRTILHDNPLSFVGN 291
QY 301 FOYLPKLTLSLNGAMDIOEPDLKGTTSLELTITLRAGIRLLPSGMCQQLRLVLELS 360
Db 292 FHNLSDLHSLVIRGASWVOQFNLGTGTHLESLTITGKISSIPNNLCOQKMLRLDLS 351
QY 361 HNOIELSLHRCOKLEETIGOHNRIMEIGADTFEQLSSQLADLSWNAIRSIHPEAFST 420
Db 352 YNNIDPESFNGCHALEELSLQRMNOITQIKETFOGLISRLDLSRNLIHEHSARFT 411
QY 421 LHSVLKLDLTDNQLTTLPLAGLGMLHLKLGNLALSOAFKSDSPKRLILEVPAYQCC 480
Db 412 LQPIITLQVDFSNELTSFTEGLNGLNQLKLVGNFKLEALAKDFVNLRLSLSPVAYQCC 471
QY 481 PYGMCASFPAKSGQWAEHLDD---EESKRPLGLLARQAEENYDQDLDELQLEMED 536
Db 472 AFWGCDSY---ANLTFENSLQDHSVAQCKTADAANTSTLENEHSOI----- 518
QY 537 SKRHPVSQCSPTPGPFKPCPEYLFESEWIRLAWAYLVLSLNGVLLTVFAGPAPLRP 596
Db 519 -----IHCPTSGAFKPCPEYLLGSMIRLTVWFIFLVALFNLVILTFPACTS-LPS 572
QY 597 VWFVVGALIGANTLIGISGLASVDALTFGQFSEYGAMWEGLGCRATGFLAVLGSEAS 656
Db 573 SKLFLGLISVSNLFGMIGITGLTFLDAVSWGRFAEFGIMWEGSGCKVAGFLAVFSES 632
QY 657 VLLTLAAVQCSVSCVAVGKSPISGIVRAGVLCGLAIGLALPLASVEGYASPL 716
Db 633 IFLMLATYERSISADIKNGKSNHLKQRYAALLAFGATYAGCFPLRHREYBASPL 692
QY 717 CLPYAPREGOPALGFTVALVMMNSFCPLVAVAGYIKLYCDLPRGFEAVWDCAMVRHVA 776
Db 693 CLPE--PTGETPSLIGFTVLLNLSLAFILMAVITKLYCNLEKEDLSSENSOSMIKHVA 750
QY 777 WLIFAGLIXCPYAFISFASMLGLPVTPEAVKSVLLVLPACINPLLYLTFNPHFR 836
Db 751 WLIFTCIFECYVAFPSFAPLITAIISIPRIMKSVTLIEFPLPACINPLVLYFNFKE 810
QY 837 D-----RRRLPRAGDSGRLAYAAAGLEK-----SSCSTQALVAFSDV 876
Db 811 DMLLKRRTYTKKSGSVSVSISSGCGLEDQFYDDCGMYSHLQGNLTVCDCCESTLLTKPV 870
QY 877 D---LILESEAGRPGLLETFFPSVTLISGQOP 907
Db 871 SCKHLI-----KSHSCPALAVASCORP 892

```

RESULT 8
AAG67556
ID AAG67556 standard; Protein; 395 AA.

XX AAG67556;
XX 26-NOV-2001 (first entry)
XX A human guanosine triphosphate binding protein coupled receptor.
XX Human guanosine triphosphate binding protein coupled receptor;
XX G protein coupled receptor; brain disease; cancer.
XX Homo sapiens.
XX WO200109323-A1.
XX 08-FEB-2001.
XX 28-JUL-2000; 2000MO-JP05070.
XX 29-JUL-1999; 99JP-0248036.

```

PR 27-AUG-1999; 99JP-0300253.  
PR 18-OCT-1999; 99US-0159590.  
PR 11-JAN-2000; 2000JP-018776.  
PR 17-FEB-2000; 2000US-0183322.  
PR 02-MAY-2000; 2000JP-0183767.  
XX  
XX (HELI-) HELIX RES INST.  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Kishimoto T;  
PI Yano K, Kanzaki K, Inoue Y;  
XX  
XX WPI: 2001-570288/64.  
XX N-PSDB; AAH78275.  
XX  
XX New gene encoding guanosine triphosphate binding protein coupled  
XX receptor, and the protein and antibodies to it, useful for diagnosis  
XX and treatment of disease such as brain disease -  
XX  
XX Claim 2; Page 48-51; 63pp; Japanese.  
XX  
XX The present sequence represents a human guanosine triphosphate binding  
XX protein coupled receptor. The guanosine triphosphate binding  
XX protein coupled receptor protein is useful in the diagnosis, prediction  
XX and treatment of disease associated with disorders of G protein coupled  
XX receptor protein, and may be useful in brain disease and cancers.  
XX  
XX Sequence 395 AA:  
SQ  
Query Match 40.1%; Score 2033; DB 22; Length 395;  
Best Local Similarity 99.5%; Pred. No. 2,7e-157;  
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 573 LLSVLNGVLLTVFAGGPRPLRPVKVVGATGANTLTGICGLASVDALTFGQFSEY 632  
Db 1 LLSVLNGVLLTVFAGGPRPLRPVKVVGATGANTLTGICGLASVDALTFGQFSEY 60  
QY 633 GARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSCVAVGKSPISGIVRAGV 692  
Db 61 GARWETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSCVAVGKSPISGIVRAGV 120  
QY 693 CLALAGLAAALPLASVEGYASPLCLPYAPREGOPALGFTVALVMMNSFCPLVAVAG 752  
Db 121 CLALAGLAAALPLASVEGYASPLCLPYAPREGOPALGFTVALVMMNSFCPLVAVAG 180  
QY 753 KLYCDLPRGFEAVWDCAMVRHVAWILFADGILYCPAFLSFASMLGLPVTPEAVKSV 812  
Db 181 KLYCDLPRGFEAVWDCAMVRHVAWILFADGILYCPAFLSFASMLGLPVTPEAVKSV 240  
QY 813 LVVLPACINPLLYLTFNPHFRDLRLRPRAGDSGRLAYAAAGLEKSSCSTQALVA 872  
Db 241 LVVLPACINPLLYLTFNPHFRDLRLRPRAGDSGRLAYAAAGLEKSSCSTQALVA 300  
QY 873 FSDVDLILESEAGRPGLLETFFPSVTLISGQOPGAPRLGSGCHVPEGNHFNQPSM 932  
Db 301 FSDVDLILESEAGRPGLLETFFPSVTLISGQOPGAPRLGSGCHVPEGNHFNQPSM 360  
QY 933 DGEILLRAEGSTPAGGGLSGGGFQPSGLAFASHV 967  
Db 361 DGEILLRAEGSTPAGGGLSGGGFQPSGLAFASHV 395  


```

RESULT 9
AAW93904
ID AAW93904 standard; Protein; 949 AA.

XX AAW93904;
XX 29-JUN-1999 (first entry)
XX Human AOMF05 protein.
XX AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain;

KW anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;
 KW diabetes; endocrine system; skeletal muscle; spinal cord; placenta;
 KW transgenic animal.

XX Homo sapiens.

XX WO915545-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US20101.

XX 24-SEP-1997; 97US-0059868.

XX (MERI) MERCK & CO INC.

XX Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;

XX WPI: 1999-254689/21.

XX Human G-protein coupled glycoprotein hormone receptor AOMF05

XX Claim 2; Fig 2; 89pp; English.

CC This invention describes the isolation of a novel human G-protein coupled
 CC glycoprotein hormone receptor, AOMF05 which has anti-obesity and
 CC anti-diabetic activity. The AOMF05 gene, receptor and modulators are
 CC useful for treatment of pancreatic disease, obesity and diabetes.
 CC Glycoprotein hormone receptors are important in the endocrine system
 CC and AOMF05 may be involved in development and function of the skeletal
 CC muscle, spinal cord, placenta and to a lesser extent, the brain. The
 CC transgenic animal may be useful for studying tissue and temporal
 CC specific expression or activity of the AOMF05 receptor, as well as
 CC for studying the ability of a variety of compounds to act as modulators
 CC of AOMF05 receptor activity.

CC Sequence 949 AA;

Query Match 40.1%; Score 2030.5; DB 20; Length 949;

Best Local Similarity 46.2%; Pred. No. 1.6e-156;

Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

QY 1 MPSPGLRALMTCALCASRRAGAPQPGPTACAPACOCEDGIMLSADSELSAV 60
 DB 1 MPPGLGL----LC--FLALGLGSGAPPGCAPPLCAAPSCGCD--RRVDSGKGLTAV 51
 QY 61 PGDDPLTAYLDLSNNLTETQPGLEHFLRLEELRLSGNHLSHIPGQAFSGLYSKIM 120
 DB 52 PEGISAFQALDISNNITQLPEDAFKPPLEELQIAGNDLSFIHPRKALSGKELKVT 111
 QY 121 LONNOLGIPALMELPISLOSLRDANLISVPERSEGLSLRLHMLDNLTEIPVR 180
 DB 112 LONNOLKTVPSAIRGSLDLSLRDANHTISVPDSFEGVLQHLMLDNLTEIPVR 171
 QY 181 ALNNLPALQAMTLALNRISHPDYAFONLTSLVVHLNNRHOHGTSEEGSLHLEETD 240
 DB 172 PLSNPPTQALTLANKLSSIPDAFTMLSSLVVHLNNKTRISQCFDGLDNLTEIPD 231
 QY 241 LNYNKLQEPVAIKTLGRLQELGFHNNNKAIPKAFMGNPLQTHFYDNPICQVGRSA 300
 DB 232 LNYNKLQEPVAIKTLGRLQELGFHNNNKAIPKAFMGNPLQTHFYDNPICQVGRSA 291
 QY 301 FQYLPKLTSLNGLAMQDQEPDLKGTSLTLITRAGIRLBSGMCQQLRRLVLELS 360
 DB 292 FHNLSDSLHSLVIRGASWQOPNLTGLVTLTGTAKISSIPNNLCQEQRMALTEIPD 351
 QY 361 HNOIEELPSLRCKLEIRIGLOHRIWEIGADTFESQSLDLSALDLSNMAISHPHAST 420
 DB 352 YNNIRDLPSFNGCHALEISIQRNQIYIKRGITQGLSILDLVSRLHIEHISRAAT 411
 QY 421 LHSIVKLDLDNOUTTLPPLAGLGLMLKLGNALSGAFSKDSFPKLILEVPAVQCC 480
 DB 412 LGRITNDVSNELTSPTEGLNGLNOLKLVGNFKLAKFVNLRSISVPAVQCC 471

QY 481 PYGMCASFPAASGOWEADLHLDDESSKRPGLLAQENHNDQDDELQLEMEDSKPH 540
 DB 472 AFMGCDSDYANLNTENNLSLODSVAQKEGTADAAWNTSTLENEHSQI----- 518
 QY 541 PSVOCSEPTPGPFKCEYLFPSWGIKRLAWVALVLLSVGNGIVLITVAGGAPAPLPYKVF 600
 DB 519 -IIHCTPSTGAFKCEYLLSSMIRLTVNFIYVALLFNLVITTTASCTG-LPSSKLF 576
 QY 601 VGAIAGANLTGISCGLASVDALTFQFSEYGARWETGLGCRATGFLAVGSEASVLL 660
 DB 577 IGLISVSNLMFGITGTILTFLDVSWGRFAEFQIMWETGSGCKVAGFLAVFSSSAIFLL 636
 QY 661 TLAAGCSVSVSCVRATKGRSPSLGVSAGVGLGCLALAGLAALPLASVGEYASPLCLPT 720
 DB 637 MLATVERSLAKDINKGNHNLKQFVALLFLGATVAGCPPLFRGYSASPLCLPT 696
 QY 721 APPEGPPALGFTVALVMNSFCFLVYAGVYIKLQCDLPBGDEAVWDCAMVHVAMLI 780
 DB 697 --PTGETPSLGTIVTLVNLNSLAFILMAVITTYTLKYLEKEDLSSENSQSMIKHVAMLI 754
 QY 781 ADGLLYCPVAFLSPASMLGLEPVTPEAVSVLLVLPACLNPLDYLLENPHRDD--- 837
 DB 755 TNCIFCPVAFESFADLITATISISPEIMKSVTLIFPLPACLNPLVYFENPKFEDWKL 814
 QY 838 -LRLRPRAAGDSGPLVYAAAGLEK-----SSCDSTQALVAFSDVD--- 877
 DB 815 LKRRVTKSGSVSVSISGQGLEDPYVDCGWYSHLQNLTVDCCESPLTRKPVSCRH 874
 QY 878 LILEASEAGRPGLTETGPPSVTLISCOOP 907
 DB 875 LI-----KSHSCPALAVASCQRP 892

RESULT 10

AAW93905

AAW93905 standard; Protein; 949 AA.

AAW93905; 29-JUN-1999 (first entry)

XX Human AOMF05 protein.

XX AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain;

XX anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;

XX diabetes; endocrine system; skeletal muscle; spinal cord; placenta;

XX transgenic animal.

XX Homo sapiens.

XX WO915545-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US20101.

XX 24-SEP-1997; 97US-0059868.

XX (MERI) MERCK & CO INC.

XX Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;

XX WPI: 1999-254689/21.

XX N-PSDB; AAX24012.

XX Human G-protein coupled glycoprotein hormone receptor AOMF05

XX Claim 1; Fig 3A-F; 89pp; English.

XX This invention describes the isolation of a novel human G-protein coupled

XX glycoprotein hormone receptor, AOMF05 which has anti-obesity and

XX anti-diabetic activity. The AOMF05 gene, receptor and modulators are

CC useful for treatment of pancreatic disease, obesity and diabetes.
 CC Glycoprotein hormone receptors are important in the endocrine system
 CC and AOMF05 may be involved in development and function of the skeletal
 CC muscle, spinal cord, placenta and to a lesser extent, the brain. The
 CC transgenic animal may be useful for studying tissue and temporal
 CC specific expression or activity of the AOMF05 receptor, as well as
 CC for studying the ability of a variety of compounds to act as modulators
 CC of AOMF05 receptor activity.

XX Sequence 949 AA;

Query Match 40.1%; Score 2030.5; DB 20; Length 949;
 Best Local Similarity 46.2%; Pred. No. 1.6e-156;
 Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

```

OY 1 MSPSPGLALMICALCALCARRAGAPOPGPPTACPAHCQEDGIMLSADCELSGLAV 60
    |||||
DB 1 MEGPLGL-----LC--FLALGLGSGAPSGAPPLCAAPCSCDGD---RRVDCSGKGLTAV 51
OY 61 PGDDPLTAYLDLSMNNLTLELPGLFHHRLFELEIRLSGNHLSHPGQAFSGLYSKITM 120
    |||||
DB 52 PEGLSAFQALDISMNNITQLPEDAFKNPFLEELQLAGNDLSFTHPKALSGLEKLVLT 111
OY 121 LONNOLGTPAALMELSLQSRDANLISVPERSEGLSLRHLMDNATLEIPVR 180
    |||||
DB 112 LONNOLKTVPSFAIRGLSALQSLRDANHTSVPEDESEGLVQRLHMDNDSLEFVPH 171
OY 181 ALNNLPALQAMTLALNRISHIPDYAFQNLTLVHLHNNRIQHLGTHSEGLHNLLETLD 240
    |||||
DB 172 PLSNPLTQALTLALNKISSIPDFATNLSLVHLHNNKIRLSQHCDFDLMLLETILD 231
OY 241 LWNKLOEFPAIRTLGLQLEGFHNNNTKAIPKAFKGNPLTQTHYDNPLOFVGRSA 300
    |||||
DB 232 LWNKNGEFPQAIKALPSLKEGFHSNSISVIPDAGFQGNPLRTIHLIDNLSLVGNSA 291
OY 301 FQYLPKRLHSLNGAMDIOEPDLKGTSTLETTLRAGIRLPSGMCQQLRNLVLELS 360
    |||||
DB 292 FHNLSLHSLVIRGASMOVQFPLTGTVLESLTILGTISSIPNNLCQEQKMLRTIDLS 351
OY 361 HNOIEELPSLHRCQKLEIGLQHNRIWEIGADTFQSOLSLQALDISWNAIRSIHEAFST 420
    |||||
DB 352 YNNIRPLSPNGCHALEEISLQNRQIYQIEGFGGLISLRIDLSRNLIHHSRAFYNT 411
OY 421 LHSVLVLDLTNOLTLPLAGLGGMLHKLKGLALSOAFKSDSPKRLILEPVAAYQCC 480
    |||||
DB 412 LGRITMLDVSENFELTFPEEGMLNQLKLVGNFKLKEALAKDFVNLISLSPVAYQCC 471
OY 481 PYGKASFPAASGOWEADLHDEESSKRPGLGLARQAEHNYDODLDELQLEMEDSKPH 540
    |||||
DB 472 AFNGCDSYALNNTENNLSIDHSAQEKGTADANVTSTLENEBSQI----- 518
OY 541 PSVOCSETPGPPEKPCERYLFESWGRILAVNAVILVSLVCNGLVLLVYFAGGPAFLPVKPY 600
    |||||
DB 519 -IHCPTSTGAFKPCETELGSMNIRLTVWFIFVALFENLITLTFASCT--LPESKLP 576
OY 601 VGMAGANTLTGSCGLASVDALTFQSEYARWETGICGRATFLAVLGSEASVLL 660
    |||||
DB 577 IGLISVNLWGLITGTILTLDAVSWGRFAEPGIMWETSGCGVYAGLAFSSESAIFLL 636
OY 661 TLAAVQCSVSCVRAVYKSPSLGSRVAGVLCIALAALPLASVGEYASPLCLPY 720
    |||||
DB 637 MLAVVERSLSKQIMKNGKSNHLKQFRVAAALFLGATVAGCFPLFRGEYSASPLCLPF 696
OY 721 APPGQGAALGFYALVWMSFCFLVYAGAVIKYCDLPBGDFEAVVDCAMVRVAVLIF 780
    |||||
DB 697 --PGETPSLGFYTLVLLNSLAEFLMAVITYTKLYCNLEKEDSENSOSSNIKRVAMLIF 754
OY 781 ADGLIYCPVAFPSMGLFEPVPEAVSVLTVLPLACLNLPLLYLFPNPHFRD-- 837
    |||||
DB 755 TNCIFPCVAFPSFAPLITATISIPDLIMKSTVTLFFPLPACLNVLVYVFEFKPKEDMKL 814
OY 838 -LRRLRPAGDSPLAVAAAGELEK-----SSCDSTQALVAFDVPD-- 877
    |||||

```

DB 815 LKRRVTKKSGSVSISQGGCLEDFYDCGMESHLOGNLTVCDCBSFLLTKPEVSCNH 874

OY 878 LTIASEAGRPPLETRYGPPSYTILSCQOP 907

DB 875 LI-----KSHSCPALAVASCORP 892

RESULT 11

AAW93965 ID AAW93965 standard; Protein; 951 AA.

XX AAW93965;

XX 29-JUN-1999 (first entry)

XX Human AOMF05 protein.

XX AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain;

XX anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;

XX diabetes; endocrine system; skeletal muscle; spinal cord; placenta;

XX transgenic animal.

XX Homo sapiens.

XX WO9915545-A1.

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US20101.

XX 24-SEP-1997; 97US-0059868.

XX (MERI) MERCK & CO INC.

XX Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;

XX WPL: 1999-254689/21.

XX Human G-protein coupled glycoprotein hormone receptor AOMF05

XX Disclosure; Fig 5; 89pp; English.

This invention describes the isolation of a novel human G-protein coupled glycoprotein hormone receptor, AOMF05 which has anti-obesity and anti-diabetic activity. The AOMF05 gene, receptor and modulators are useful for treatment of pancreatic disease, obesity and diabetes. Glycoprotein hormone receptors are important in the endocrine system and AOMF05 may be involved in development and function of the skeletal muscle, spinal cord, placenta and to a lesser extent, the brain. The transgenic animal may be useful for studying tissue and temporal specific expression or activity of the AOMF05 receptor, as well as for studying the ability of a variety of compounds to act as modulators of AOMF05 receptor activity.

XX Sequence 951 AA;

Query Match 40.1%; Score 2030.5; DB 20; Length 951;
 Best Local Similarity 46.2%; Pred. No. 1.6e-156;
 Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

```

OY 1 MSPSPGLALMICALCALCARRAGAPOPGPPTACPAHCQEDGIMLSADCELSGLAV 60
    |||||
DB 1 MEGPLGL-----LC--FLALGLGSGAPSGAPPLCAAPCSCDGD---RRVDCSGKGLTAV 51
OY 61 PGDDPLTAYLDLSMNNLTLELPGLFHHRLFELEIRLSGNHLSHPGQAFSGLYSKITM 120
    |||||
DB 52 PEGLSAFQALDISMNNITQLPEDAFKNPFLEELQLAGNDLSFTHPKALSGLEKLVLT 111
OY 121 LONNOLGTPAALMELSLQSRDANLISVPERSEGLSLRHLMDNATLEIPVR 180
    |||||
DB 112 LONNOLKTVPSFAIRGLSALQSLRDANHTSVPEDESEGLVQRLHMDNDSLEFVPH 171
OY 181 ALNNLPALQAMTLALNRISHIPDYAFQNLTLVHLHNNRIQHLGTHSEGLHNLLETLD 240
    |||||

```

```

Db 172 PLSNLPFLQALTLALNKISSIPDFAFTNLSLVYHLHNNKISLSQHCDFGLDNETLD 231
QY 241 LNNKXIOEFFVAIRTRIGLOELCFHNNNNKAIPEKAFMGNDLQTHFYDNPLOFVGSA 300
Db 232 LNNNGEPPQAIKALPSIKELGPHNSNISVDPGAFDGNPLRTIHLNDNPISFVNSA 291
QY 301 FOYLPKLTLSLNGAMDIQEFPLDKGTTSEILTLTRAGIRLLPSGCOQLPRLVLELS 360
Db 292 FHNLSDLHSLVIRGASWVOQFPMLTGVHLESLITLGTKISSIPNNLCQEQKMLRTLDLS 351
QY 361 HNOIEPLSLHRCQKLEELGLOHNRIMEIGADTFQSLSQALDLSMNAIRSIHPEAFST 420
Db 352 YNNIRDLPSFNGCHALEEISIQRNQIYQIEGTFQGLISIRLIDVSRNLIHEHSRAFAT 411
QY 421 LHSIVKLDLTDNQLTTLPLAGLGMHLKGNLALSOAFSKDSFPKRLILEVYAYQCC 480
Db 412 LGPITNDVSPFNELTSFPEGLNGLNQKLVGNFKELKEALADFNVLRLSLSPYAYQCC 471
QY 481 PYGCASFEEKASGQWAEEDLHLDDEESSKRPGLGLARQAEHNDODLDELQLEMEDSKPH 540
Db 472 AFWGCDSYANLNTENNISLODHSVAQEKGTADAANVTSTLENEHSQI----- 518
QY 541 PSVQCSPTPGPFKPCBYLEESGIRLAWAYLVLSYLCGLVLLTVFAGGAPLPVPKVF 600
Db 519 -IIHCTPSTGAFKPCBYLLGSMIRLTVWFIFLVALFENLVLITLTFASCTSLPSKLE 576
QY 601 VQAGANTVTGISCGLASVDALTFQGFSEYGARWETGLGRATGFLAVLSSEAVLL 660
Db 577 IGLISVSNLMFMGIYGLITFLDAVSMGRFAEFGIMWETSGCKVAGFLAVSESEAFIL 636
QY 661 TLAAGVGSVSCVAVAKSPSLGVRAGVLCGLALAGLAAALPLASVGEYASPLCLPY 720
Db 637 MLATYERLSAKADIKKNSKSHLQKFRVAAALALFAGATVAGGCFPLFHGEYSASPLCLPF 696
QY 721 APPEGQPALGETVALVAMNSFCFLVAGAYIKIKCDLPRGFEAVWMCAMRHANMLIF 780
Db 697 --PGEFSPISGFTVTLVNLNLAFLMAVITKLCNLEKEDLSNOSSMKHNAMLIF 754
QY 781 ADGLXCGVAFSLFASMLGLFPVPEAVKSVLVVLPPLACINPLLYLLEFPHFRD--- 837
Db 755 TNCIFECVAFSFAPLITTAISIPETIKSVTLIFPLPACINPLVLYEFNPKFEDMKL 814
QY 838 -LRLRPRAAGSGPLAYAAAGELK-----SSQDSQALAFSPVD--- 877
Db 815 LKRVRTKKSGSVSISISQGCLEODPYDQGMSHLDGNLFLVCCDCESFLLTKRVSCN 874
QY 878 LILEASERAPPGLETYGFPSVTLISQOP 907
Db 875 LI-----KSHSCPALAVASCORP 892

```

RESULT 12

AAW93906
ID AAW93906 standard; Protein; 951 AA.

AC AAW93906;

DT 29-JUN-1999 (first entry)

DE Human AOMF05 protein.

KM AOMF05; human; G-protein coupled glycoprotein hormone receptor; brain;
anti-obesity; anti-diabetic; treatment; pancreatic disease; obesity;
diabetes; endocrine system; skeletal muscle; spinal cord; placenta;
transgenic animal.

OS Homo sapiens.

XX W09915545-A1.

PD 01-APR-1999.

XX

```

PF 24-SEP-1998; 98MO-US20101.
XX
PR 24-SEP-1997; 97US-0059868.
XX
PA (MERK) MERCK & CO INC.
XX
PI Abramovitz M, Liu Q, McDonald TP, O'Neill GP, Wang R;
XX WPI; 1999-254689/21.
XX N-PSDB; AAX24014.
XX
PS Human G-protein coupled glycoprotein hormone receptor AOMF05
XX disclosure; Fig 6A-F; 89pp; English.
XX
CC This invention describes the isolation of a novel human G-protein coupled
CC glycoprotein hormone receptor, AOMF05 which has anti-obesity and
CC anti-diabetic activity. The AOMF05 gene, receptor and modulators are
CC useful for treatment of pancreatic disease, obesity and diabetes.
CC Glycoprotein hormone receptors are important in the endocrine system
CC and AOMF05 may be involved in development and function of the skeletal
CC muscle, spinal cord, placenta and to a lesser extent, the brain. The
CC transgenic animal may be useful for studying tissue and temporal
CC specific expression or activity of the AOMF05 receptor, as well as
CC for studying the ability of a variety of compounds to act as modulators
CC of AOMF05 receptor activity.
CC
XX
SQ Sequence 951 AA:

```

```

Query Match 40.1%; Score 2030.5; DB 20; Length 951;
Best Local Similarity 46.2%; Pred. No. 1.6e-156;
Matches 430; Conservative 133; Mismatches 306; Indels 61; Gaps 10;

```

```

QY 1 MSPPGALMICAALCASRRAGAPQPGPTACPARCHCEDIMLSADCELSGLSAV 60
Db 1 MEGPLGL----LC--FLALIGLSAGPSGAAPRLCAAPCCDCG--RRVDCSGGLTAV 51
QY 61 PGDDPLTAYIDLMSNNLTLOPLGFHHLRFLTEELRSGNHSHIRPGAFSGLSKLTM 120
Db 52 PEGLSAFQALDISNNITQLPEDAFKFPPLLEEQQLGNDLSFTHRPALSGELKLYLT 111
QY 121 LQNNQGGIPAEALWELPSLOSLRLDANLISLVERSEFGLSSLRHLWDNNALEIVR 180
Db 112 LQNNQKTVPESEAIRGSLQSLRLDANHITSVPDESFGYQLRHMLDQNSLIEVYH 171
QY 181 ALNNLPALQAMTLALNRTSHIDYAFQNLTSVLYLHLNNKIOHLGHSFEGLANLETLD 240
Db 172 PLSNLPFLQALTLALNKISSIPDFAFTNLSLVYHLHNNKISLSQHCDFGLDNETLD 231
QY 241 LNNKXIOEFFVAIRTRIGLOELCFHNNNNKAIPEKAFMGNDLQTHFYDNPLOFVGSA 300
Db 232 LNNNGEPPQAIKALPSIKELGPHNSNISVDPGAFDGNPLRTIHLNDNPISFVNSA 291
QY 301 FOYLPKLTLSLNGAMDIQEFPLDKGTTSEILTLTRAGIRLLPSGCOQLPRLVLELS 360
Db 292 FHNLSDLHSLVIRGASWVOQFPMLTGVHLESLITLGTKISSIPNNLCQEQKMLRTLDLS 351
QY 361 HNOIEPLSLHRCQKLEELGLOHNRIMEIGADTFQSLSQALDLSMNAIRSIHPEAFST 420
Db 352 YNNIRDLPSFNGCHALEEISIQRNQIYQIEGTFQGLISIRLIDVSRNLIHEHSRAFAT 411
QY 421 LHSIVKLDLTDNQLTTLPLAGLGMHLKGNLALSOAFSKDSFPKRLILEVYAYQCC 480
Db 412 LGPITNDVSPFNELTSFPEGLNGLNQKLVGNFKELKEALADFNVLRLSLSPYAYQCC 471
QY 481 PYGCASFEEKASGQWAEEDLHLDDEESSKRPGLGLARQAEHNDODLDELQLEMEDSKPH 540
Db 472 AFWGCDSYANLNTENNISLODHSVAQEKGTADAANVTSTLENEHSQI----- 518
QY 541 PSVQCSPTPGPFKPCBYLEESGIRLAWAYLVLSYLCGLVLLTVFAGGAPLPVPKVF 600
Db 519 -IIHCTPSTGAFKPCBYLLGSMIRLTVWFIFLVALFENLVLITLTFASCTSLPSKLE 576

```

therapy, and can be used as nutritional supplements. They may be used to

QY 878 LILEASEGRPGLFTYGFPSVTLISCOOP 907

XX 24-SEP-1997; 97US-0059863.
XX (MERI) MERCK & CO INC.

XX PI Bailey WJ, Liu Q, McDonald TP;
XX WPI: 1999-254711/21.

DR N-PSDB; AAX23981.

XX Human G-protein coupled glycoprotein hormone receptor HG38

XX PS Disclosure; Fig 3A-E; 74pp; English.

XX CC This invention describes a novel human G-protein coupled glycoprotein
XX CC hormone receptor, HG38. Glycoprotein hormone receptors are important in
XX CC the endocrine system and HG38 may be involved in development and function
XX CC of the skeletal muscle, spinal cord, placenta and to a lesser extent, the
XX CC brain. The transgenic animal may be useful for studying tissue and
XX CC temporal specific expression or activity of the HG38 receptor, as well as
XX CC for studying the ability of a variety of compounds to act as modulators
XX CC of HG38 receptor activity.

XX SQ Sequence 692 AA;

Query Match 39.4%; Score 1996; DB 20; Length 692;

Best Local Similarity 57.9%; Pred. No. 6,7e-154;

Matches 393; Conservative 95; Mismatches 171; Indels 20; Gaps 6;

OY 23 GGAPQPGPPTACAPACHOEDG-IMLSADCSRLGSAVPGDLDPLTAYIDLSMNNLTSL 81
DB 22 GSSPRSGVLLRCGPTCHCEPDGRMLRYDCSDGLSELPSNLSVTYIDLSMNNISQL 81
OY 82 QPGLFHLHLFLEELRLSGNHLSHIPGQAFSGYSLKILMLQNNQGLGIPAEALMELPSLQ 141
DB 82 LNPPLPSLFLFELRLAGNALTYIPKGAFTGLYSLKVLMLQNNQGLRHVPTEALQNLRSIQ 141
OY 142 SLRLDANLISLPERFEGLSIRHMLDNDNATEIPVALNNTLPALQAMTLALNRISHI 201
DB 142 SLRLDANLISLPERFEGLSIRHMLDNDNATEIPVALNNTLPALQAMTLALNRISHI 201
OY 202 PDYAFONTLSIVYLHNNRRIHGLTHSEFGLHNLLETLDLNNKLOEFPVARTLGRLOE 261
DB 202 PDYAFONTLSIVYLHNNRRIHGLTHSEFGLHNLLETLDLNNKLOEFPVARTLGRLOE 261
OY 262 LGFHNNNIKAIPKAFWGNPLQTIHFYDNPLOFVGRSAFOYLPKLTLSLNGAMDIOEF 321
DB 262 LGFHNNNIKAIPKAFWGNPLQTIHFYDNPLOFVGRSAFOYLPKLTLSLNGAMDIOEF 321
OY 322 PDLAGTTSLEILLTRAGIRLLPSGMOQOLPRLRVLELSHNOIEELPSLRCKLEIGL 381
DB 322 PDLAGTTSLEILLTRAGIRLLPSGMOQOLPRLRVLELSHNOIEELPSLRCKLEIGL 381
OY 382 QHNRIMEIGADTFESQSLQALDLSMNAISIHPEAFSTLSLVKIDLTDNQLTTLPLAG 441
DB 382 QHNRIMEIGADTFESQSLQALDLSMNAISIHPEAFSTLSLVKIDLTDNQLTTLPLAG 441
OY 442 LGLMLHLKLGKALALSOAFKSKDSFPKRLILEVPRAYOCCPYGMCASFVKASGOW----- 495
DB 442 LGLMLHLKLGKALALSOAFKSKDSFPKRLILEVPRAYOCCPYGMCASFVKASGOW----- 495
OY 496 EAEIDLHDESSSRPRLGLLARAQENHYDDDELQLEM-EDSKRHPSPVOCSPTPGPFPK 554
DB 496 EAEIDLHDESSSRPRLGLLARAQENHYDDDELQLEM-EDSKRHPSPVOCSPTPGPFPK 554
OY 502 SMDLH-----KKDAGMFOAQDE-----RLDEFLDDEFEDLKALHSVQCSPPGPFPK 550
DB 502 SMDLH-----KKDAGMFOAQDE-----RLDEFLDDEFEDLKALHSVQCSPPGPFPK 550
OY 555 CEYLFEWSGIRLAWAIVLVLVLCNGVLVLTVERAGRAPLDPVKFVVGAIGANTLTGIS 614
DB 555 CEYLFEWSGIRLAWAIVLVLVLCNGVLVLTVERAGRAPLDPVKFVVGAIGANTLTGIS 614
OY 551 CERHLDGMLIRIGVWTLAVLALTCNALVTSTVFR-SPLYISPIKILIGVIAVNNLTGVS 609
DB 551 CERHLDGMLIRIGVWTLAVLALTCNALVTSTVFR-SPLYISPIKILIGVIAVNNLTGVS 609
OY 615 CGLASVDALFFGQSEYGAWENGIGCRATGFLAVGSEASVLLTLAAVOCSSVSCV 674
DB 615 CGLASVDALFFGQSEYGAWENGIGCRATGFLAVGSEASVLLTLAAVOCSSVSCV 674
OY 610 SAVIAGVDAFTFGSFARHGAWENGVCCHVIGLSIFRASESSVFLITLALERGSVAKS 669
DB 610 SAVIAGVDAFTFGSFARHGAWENGVCCHVIGLSIFRASESSVFLITLALERGSVAKS 669

OY 675 RAYGKSPSLGSRVAGVLCG 693
DB 670 AKFTKAPPSSILKVIILLC 698

Search completed: October 23, 2002, 10:43:27
Job time : 43 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52; Search time 26 Seconds
(without alignments)
3573.784 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068
Sequence: 1 MPSPGIRALMLCALICASR.....GGLSGGGGFPSPGIAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 71:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2525.5	49.8	907	2 JE0176	orphan G protein-c
2	2443.5	48.2	907	2 JC0193	foliitropin-coupled
3	534	10.5	694	2 JC4301	foliitropin receptor
4	526.5	10.4	695	2 I45896	foliitropin receptor
5	525	10.4	694	2 JC2237	foliitropin receptor
6	524.5	10.3	695	2 JC1493	foliitropin receptor
7	519.5	10.2	692	2 A34548	foliitropin receptor
8	518	10.2	695	1 ORHUT	foliitropin receptor
9	515	10.2	695	1 JN0898	foliitropin receptor
10	490.5	9.7	764	2 A40077	thyrotropin receptor
11	486	9.6	764	2 I48882	thyrotropin receptor
12	484.5	9.6	1091	2 A58532	glial cell membran
13	476	9.4	700	2 JC7389	thyroid stimulat
14	474	9.4	700	2 I77463	luteinizing hormon
15	474	9.4	700	2 A49744	lutropin-choriogon
16	473.5	9.3	764	2 JC3643	thyroid stimulat
17	472.5	9.3	793	2 JC7390	thyroid stimulat
18	472	9.3	700	2 A42395	lutropin receptor
19	469.5	9.3	764	1 ORHUR	thyrotropin receptor
20	467	9.2	696	2 A41344	lutropin-choriogon
21	463	9.1	764	2 JC7361	foliitropin receptor
22	462.5	9.1	764	2 A35956	thyrotropin receptor
23	455.5	9.0	603	2 JC1282	insulin-like growth
24	439.5	8.7	699	1 ORHUT	lutropin-choriogon
25	439	8.7	605	2 JC5239	insulin-like growth
26	438.5	8.7	603	2 JC6128	insulin-like growth
27	437.5	8.6	605	2 A41915	insulin-like growth
28	432.5	8.5	889	2 T20123	hypothetical prote
29	411	8.1	925	2 JC2033	G protein-coupled

30	397	7.8	1469	2 B36655	silt protein 2 pre
31	397	7.8	1480	2 A36655	silt protein 1 pre
32	394.5	7.8	1531	2 T42218	silt-1 protein hom
33	384.5	7.6	1066	2 T15864	hypothetical prote
34	384	7.6	1385	2 T13887	tlr protein - frui
35	383.5	7.6	1523	2 T13953	MEGF5 protein - ra
36	382	7.5	1389	2 T13852	gene wheeler prote
37	375	7.4	1389	2 T13852	platelet membrane
38	375	7.4	1119	2 AD1822	leucine-rich-repea
39	364.5	7.2	1134	1 A29944	chaoptin precursor
40	363	7.2	536	2 A34901	lysine carboxypept
41	363	7.2	662	2 S42799	gap precursor
42	360.5	7.1	738	2 T19938	hypothetical prote
43	354.5	7.0	961	2 T23395	hypothetical prote
44	348.5	6.9	1115	2 S40241	G protein-coupled
45	340.5	6.7	1039	2 T22117	hypothetical prote

ALIGNMENTS

RESULT 1

JE0176 orphan G protein-coupled receptor precursor - human

C.Species: Homo sapiens (man)

C.Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000

C.Accession: JE0176

R.McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.

Biochem. Biophys. Res. Commun. 247, 266-270, 1998

A.Title: Identification and cloning of an orphan G protein-coupled receptor of the gl

A.Reference number: JE0176; M0ID:98508104

A.Accession: JE0176

A.Molecule type: mRNA

A.Residues: 1-907 <MCID>

A.Cross-references: GB:AF062005; NID:q3366801; PIDN:AAC28019.1; PID:q3366802

C.Comment: This protein is a receptor for a novel class of glycoprotein ligands.

C.Genetics:

A:Gene: HG38

A:Map position: 12q22-23

F:1-21/Domain: signal sequence #status predicted <SIG>

F:562-583/Domain: transmembrane #status predicted <TM1>

F:594-616/Domain: transmembrane #status predicted <TM2>

F:639-660/Domain: transmembrane #status predicted <TM3>

F:681-701/Domain: transmembrane #status predicted <TM4>

F:725-744/Domain: transmembrane #status predicted <TM5>

F:768-791/Domain: transmembrane #status predicted <TM6>

F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 49.8%; Score 2525.5; DB 2; Length 907;

Best Local Similarity 55.3%; Pred. No. 1.2e-168;

Matches 497; Conservative 136; Mismatches 235; Indels 31; Gaps 9;

QY	23	GGAPQPGPPTACPAFCOEDG-IMLSADCSSEGLSAVGGDDLPATAYLDSNNITEL	81
DB	22	GSSRSRVLLRGCTHCHCEPDGMLLRVDCSDGLSELPNSLVSFTYDLSSNNISQL	81
QY	82	QPGLEHRLRELELRSLGSHLSHPGQFSGLSLKLIMLQNNQDGLIPALMELSLQ	141
DB	82	LPNPLPSLRFLELELRSLGSHLSHPGQFSGLSLKLIMLQNNQDGLIPALMELSLQ	141
QY	142	SLRLDNLISLVSPEFSGSLSLRLHMLDDNALFEIPRALNIPALQAMTLALNRISHI	201
DB	142	SLRLDNLISLVSPEFSGSLSLRLHMLDDNALFEIPRALNIPALQAMTLALNRISHI	201
QY	202	PDVAFOMLTSLVLAHLHNNRIQLHSHSEGLANLETLIDLYNKLOEPFAIRLGRLOE	261
DB	202	PDVAFOMLTSLVLAHLHNNRIQLHSHSEGLANLETLIDLYNKLOEPFAIRLGRLOE	261
QY	262	LGFINNNIKALPEKAFKCNPLQTHFYDNPLOVYGRSAFOYLPKRLTSLSGANDIOEF	321
DB	262	LGFINNNIKALPEKAFKCNPLQTHFYDNPLOVYGRSAFOYLPKRLTSLSGANDIOEF	321
QY	322	PDLKGTSLLEITLTLTRAGIRLPLPSGMCQOLPRLRYLDELSHNOIEELPSLRQCUIEIGL	381

```

Db 322 PDLTGANTLESILTLTGAQSSILPQIVCNOLPNIQVLDLSLNLEDDLPSPVCKKQKIDL 381
Qy 382 OHNRIMEIGADTFSSQSLQALDLSWNAIRSHPEAFSTLHSLVKLDLTDNQLTTLPLAG 441
Db 382 RINEIYEIVDFQQLSLRSLNLAWNKTAIHPNAFSTPLSLIKDLSSNLSPPITG 441
Qy 442 LGLMLHLKLGKLNALSOAFSKDSFPKRLILEVPAVQCCPYGMCASFRRASGM----- 495
Db 442 LHGLTHLKLTLGNHALQSLSSSEFPELKIEMPRATQCCAFVCENAYKISNOMKGDNS 501
Qy 496 EAEEDLHLDDESSKRPGLILARQAEHNDQDLDELQLEM-EDSKPHPSVQCSPPGPFKRP 554
Db 502 SMDDDH-----KKDAGMFOAQDE---RDLEDFLDDEEDLKALHSVQCSPPGPFKRP 550
Qy 555 CEYLEFESMGIRLAWNAVATLVSLCNGVLTLYFAGAPALPVPKFEVVGAGIAGANTLTGIS 614
Db 551 CEHLIDGMLIRIGVMTIAVLTCNALVTSTVR-SPLYISPIKILLIYIAVNMILTGS 609
Qy 615 CGLLASVDALTFPGSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVOCSSVSCV 674
Db 610 SAVLAGVDAFTFGSPARHAGAMWENGCHVIGFLSTFASSEVFLTLTALERGSVKYS 669
Qy 675 RAYGSPSLGSRACVLCGLALAGLAALPLASVEYSGASPLCLPYAPPEGAPALGFTV 734
Db 670 AKETKAPFSSILKVTILLCALALATMAAVPLLGSGSKYGASPLCLPL--PFGSPSTMGWY 727
Qy 735 ALVMNSFCFLVAVAGATYIKLYCDLPRGDFEAVDCAMVRHVMILFADGLLYCPAFLSF 794
Db 728 ALIILNSICFLMTATYIKLYCNLDKGLDENTWDSNVKHIALLFTYICINPAVFLSF 787
Qy 795 ASMLGLFVTPPAVKSVLLVPLPACINPLLYLLENPHFRDLRLRPRA-----GDGSG 849
Db 788 SSLINLTPISEVFIKILLVAVPLPACINPLLYLLENPHFRDLRLRPRA-----GDGSG 849
Qy 850 PLAVAAAGELKSSDSTQALVAFSDVDLLEASAGRP---PGLETYGFPSTVLISC 904
Db 848 SLMSINSDVEKQSDSTQALVFTSSITVDLPPSVSPAYPTESCHLSVAFVPC 906

RESULT 2
G protein-coupled receptor FEX - mouse
C:Species: Mus musculus (house mouse)
C>Date: 23-Jul-1999 #sequence.revision 23-Jul-1999 #text.change 11-May-2000
C/Accession: JG0193
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A>Title: Identification of a novel seven-transmembrane receptor with homology to glycoprotein A; Accession: JG0193; MUID:99121227
A:Reference number: JG0193; MUID:99121227
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>

```

```

Query Match 48.2%; Score 2443.5; DB 2; Length 907;
Best Local Similarity 54.3%; Pred. No. 6.9e-163;
Matches 485; Conservative 146; Mismatches 244; Indels 19; Gaps 9;

```

```

Qy 22 AGGAPGPGPTACPAFCHQCEQDG-IMLSADCSFELSAVPGDDPLVATYDLSMNNITE 80
Db 21 AGSSPGDAPRPGPSCHCELDGRMLRLKRDSCSLGSELPSNLSVTSTYDLSMNNISQ 80
Qy 81 LQGLPLHFLFLEELRLSGHLSHIEQAFSGLYSKITLMQNNQGLGIPAAELWELPST 140
Db 81 LPASTLHRLCFLEELRLAGALTHIPKGAFTGHSHTKVLMLQNNQRLKVEEALQNRSL 140
Qy 141 OSRLDANLISLVERSEFELSLRHLMDNMLTEIPVALNNRPALQAMTALANISH 200
Db 141 OSRLDANLISLVERSEFELSLRHLMDNMLTEIPVALNNRPALQAMTALANISH 200
Qy 201 IDPYAFQNTLSLVVHLHNNRIQHLGTHSPEGLHNLDTLDLNNKLOEFPVATRTGLQ 260
Db 201 IDPYAFQNTLSLVVHLHNNRIQHLGTHSPEGLHNLDTLDLNNKLOEFPVATRTGLQ 260

```

```

Db 201 IDPYAFQNTLSLVVHLHNNRIHSLKCPDGLHSETLTDLNNNDEPPTAITSNLEK 260
Qy 261 ELGFHNNNIKAIPKAFMGNPLLOTITHEFYDNPIQFGRSAFOYLPKLTHTLSGANDIOE 320
Db 261 ELGFHNNNIKASIPERAFVGNPSLITTHFDNPLOFVGSAFOHLPKLTHTLSGASHITE 320
Qy 321 FPDKITSLEIITLRAGIRLLPSGMCOQLPRLRVLETSHNDIEELPSLRCKLEETG 380
Db 321 FPHLTGATLELTLTGAKISSLPQAVCDOLPNIQVLDLSYNLEDDLPSSGCKKQKIDL 380
Qy 381 LOHNRIMEIGADTFSSQSLQALDLSWNAIRSHPEAFSTLHSLVKLDLTDNQLTTLPLA 440
Db 381 RINEIYEIVDFQQLSLRSLNLAWNKTAIHPNAFSTPLSLIKDLSSNLSPPITG 440
Qy 441 LGLMLHLKLGKLNALSOAFSKDSFPKRLILEVPAVQCCPYGMCASFRRASGM----- 500
Db 441 GLHGLTHLKLTLGNHALQSLSSSEFPELKIEMPRATQCCAFVCENAYKISNOMKGDNS 500
Qy 501 HLDDESSKRPGLILARQAEHNDQDLDELQLEM-EDSKPHPSVQCSPPGPFKPEYLF 559
Db 501 NSVD-DLHKKDAGLRFQVODE---RDLEDFLDDEEDLKALHSVQCSPPGPFKPEYLF 555
Qy 560 ESMGIRLAWNAVATLVSLCNGVLTLYFAGAPALPVPKFEVVGAGIAGANTLTGISGLLA 619
Db 556 GSWMLRIGVMTIAVLTCNALVATLVR-TPLYISIKILLIYIAVNDIIMGVSSAVIA 614
Qy 620 SYDALTFQGFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVOCSSVSCVAYK 679
Db 615 AYDAFTFGFRAHAGAMWENGCHVIGFLSTFASSEVFLTLTALERGSVKSKSEY 674
Qy 680 SPSLGSVAGVLCGLALAGLAALPLASVEYSGASPLCLPYAPPEGAPALGFTV 739
Db 675 KAPLSTLRATVLCVLLALTLTIPDLGSKYNASPLCLPL--PGEESTGYMALVLL 732
Qy 740 NSFCELVAVAGATYIKLYCDLPRGDFEAVDCAMVRHVMILFADGLLYCPAFLSF 799
Db 733 NSICPLMTIATYIKLYCSEKGELENMWDNSVKHIALLEFANCLILYCPAFLSFSSLIN 792
Qy 800 LEPTPPEAVKSVLLVPLPACINPLLYLLENPHFRDLRLRPRA-----GDGSPPLA 854
Db 793 LFTSPDVIKELVILVPLPACINPLLYLLENPHFRDLRLRPRA-----GDGSPPLA 852
Qy 855 AAGELEKSSDSTQALVAFSDVDLLEASAGRP---PGLETYGFPSTVLISC 904
Db 853 NSDVEKRSCESTQALVFTASTIAYDLPSTGASPAVPTESCHLSVAFVPC 906

```

RESULT 3

JC4301

folliotropin receptor - pig

N:Alternate names: follicle-stimulating hormone receptor

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 16-Nov-1995 #sequence.revision 08-Feb-1996 #text.change 21-Jan-2000

C/Accession: JC4301

R:Remy, J.-J.; Labib-Manssals, Y.; Yerville, M.; Bozon, V.; Couture, L.; Pajot, E.; Grebe

Gene 163, 257-261, 1995

A>Title: The porcine follicliotropin receptor: cDNA cloning, functional expression and chr

A:Reference number: JC4301; MUID:96011644

A:Accession: JC4301

A:Molecule type: mRNA

A:Residues: 1-694 <REM>

A/Cross-references: GB:L31966

A/Experimental source: ovarian granulosa cells

A/Comment: This receptor belongs to the family of the G-protein coupled receptors. It

encodes a protein in male and oogenesis in female.

A/Genetics:

A:Gene: fshr

A:Map position: 3 q2.2-q2.3

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repa

C/Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein

F:1-365/Domain: follicle-stimulating hormone binding #status predicted <IOB>

F:170-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

F:366-388/Domain: transmembrane #status predicted <TM1>

A;Gene: FSHR

F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F;56-70/Domain:	leucine-rich alpha-2-glycoprotein repeat	homology	<LRK1>
F;71-85/Domain:	leucine-rich alpha-2-glycoprotein repeat	homology	<LRK2>

F;56-70/Domain:	leucine-rich alpha-2-glycoprotein repeat	homology	<LRK1>
F;71-85/Domain:	leucine-rich alpha-2-glycoprotein repeat	homology	<LRK2>

Query Match	10.4%	Score 525;	DB 2;	Length 694;
Best Local Similarity	23.8%	Pred. NO. 1, 1e-28;		
Matches	185;	Conservative 122;	Mismatches 305;	Indels 164;
				Gaps 21;

RESULT 6
JC1493
follicleotropin receptor - sheep

Query Match	10.38;	Score 524.5;	DB 2;	Length 695;
Best Local Similarity	26.4%;	Pred. No. 1.2e-28;		
Matches 175;	Conservative 101;	Mismatches 273;	Indels 113;	Gaps 18

```

QY 261 ELGHHNNNRKIKPEKAFNGNLLQTHFYDNI-QFVGSAAQYLPKHTLTLINGAMDIO 319
Db 50 ELKRVTLKRVKIPBGASGFEDLEKITSQNDVLEVIENAVNSNLPKEHRIERIEKANNTLL 109
QY 320 EF-PDLKGTTSLELITLTLRAGIRLLPSGMCODPLRYVLELHNOJEEPLSHROCKLEE 378
Db 110 YIDPD-----AFQNLPRMLRILLINTGTGKHLPAVHKIOSLQK 146
QY 379 I--GLQHN-RIMEGATFQSLS-SLOALDISMNAIRSTIHPEAF-----STL 421
Db 147 VLLDIOQNIHITVERNSFMGLSESMITWVLSKNGIOETHNCAFNGTQDLELNLSDNSNL 206
QY 422 HSL-----VKDLDTNOLTTPLPLAIGGLMHKIKGNLALSQAFNSDPKRLI 470
Db 207 EELPNVYFQGASGVILIDISKTRISLPYIGLENLKLKRASTYHLLKPLSLEKVTYWE 266
QY 471 LEVPYVYQCCPYGMCASFRRASGGWEADLH-----LDDEESSKRPJGLIARQA 519
Db 267 ASLTYPSHCCAF-----ANMRQTSDLHPICNKSILRQEVDMQARGRISLAEDD 318
QY 520 ENHYDDDLDELQLEMEDS--KPHRSVCCSPRPGRPKCEYULEPSKGRILAVAWAILSVL 577
Db 319 EBSYAGFDMMSSEFYDILDCEVYDVYTCSEPRDAFNCEDIMGYDILRVLIWFIITLAIIT 378
QY 578 CNGILVLYEYFAGGPAFLPYKFFVYAGTAGANTLIGISGLIASDALTLFGQFSEYGAEME 637
Db 379 GNLIVLV-ILITSQKYLTVRPFILMCNIAFADLCIGITVLLILASVDVHKKSOYHNNAIDMO 437
QY 638 TELGCRATGFLAVLGESEASYLLTLTAVQ-----CSVSVC--VRAYGKSPSISGVRA 688
Db 438 TGAGCDMAAGFFVYFASSELASYTLTALTLERWHTITTHMOJECQKVNHNHNASIMLVGWV-- 495
QY 689 GVLGCLALAGLAAALPLASVGEYGASPLCLPYAPBEGQAPALGFTVALVMNNSFCFLVYA 748
Db 486 -----FAFAVALPFTGISSTYMKVSGICLP--DIDSPSOLTYVMSLLVLTVALFAFYIC 546
QY 749 GAYIKYLCDLPRGDF-EAWWDCAMVRHVAVLIPADGLLYCPVAFLPASAISGLFPRPPEA 807
Db 547 GCYTHILTVLRNPNTITSSSDTKIARKNAMLITTFELCMAPISFPAISASIKVPLITVSK 606
QY 808 VKSVLLVVLPLPACLNPLLYLLFNPHFRLDLRLRPRAGDSGPLVAAAGELEKSSCDST 867

```

Db 607 SKILLVLEFPLNSCANFLVAIFTRNFRDRDFIL-----LSKFGCEYV 649
 QY 868 QA 869
 Db 650 QA 651

RESULT 7

A34348
 Folliotropin receptor precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
 C:Accession: A34548; A41729
 R:Spengler, R.; Braun, T.; Nikolics, K.; Segaloff, D.L.; Seeburg, P.H.
 Mol. Endocrinol. 4, 525-530, 1990
 A:Title: The testicular receptor for follicle stimulating hormone: structure and function
 A:Reference number: A34548; M01D:91125358
 A:Accession: A34548
 A:Residues: 1-692 <SP>
 A:Molecule type: mRNA
 A:Cross-references: GB:102842; M1D:9204183; P1D:AAA41175.1; P1D:9204184
 R:Heckert, L.L.; Daley, J.J.; Griswold, M.D.
 Mol. Endocrinol. 6, 70-80, 1992
 A:Title: Structural organization of the follicle-stimulating hormone receptor gene.
 A:Reference number: A41729; M01D:92149579
 A:Accession: A41729
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-692 <HEC>
 A:Cross-references: GB:S81198; M1D:9245344; P1D:AA821415.1; P1D:9245345
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:81117, NCBIN:81119, NCBIN:81121, NC
 Mol. Endocrinol. 9, 159-170, 1995
 A:Title: Identification of the sites of N-linked glycosylation on the follicle-stimulating
 A:Reference number: A57562; M01D:95295729
 A:Contents: annotation; glycosylation sites
 C:Function:
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-692/Domain: follicle-stimulating hormone binding #status predicted <MAT>
 F:16-366/Domain: extracellular hormone binding #status predicted <EHB>
 F:56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:367-387/Domain: transmembrane #status predicted <TM1>
 F:398-421/Domain: transmembrane #status predicted <TM2>
 F:443-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191,199,293/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Db 82 VLEVIDADYPSNLPKLHEIRIEKANMLTYINPEAFONPLSLRYLLISNTGKHL-----P 136
 QY 232 GLHNLETLIDNINLKOEPPVAIRTLGRLOELGFHNNNIKAIEKAPMGNPLQTHIFYDN 291
 Db 137 AVHKIOSIQ-----KVLIDID-----NINITHVANRSMGMS-----FES 172
 QY 292 PLOFVGRSAFOYLPKLTHTLSANGAMDIOEPDLKGTTSLEITLTPRAGIRLLPSGMCOL 351
 Db 173 VILMLSKNGIE-----EIHNCAPNG-----TQDELINLS----- 201
 QY 352 PRLVLELSHNOIEPLPSLHRCOKLEIFIGLQHRNIEIGADPFSSQSLQALDSNNAIR 411
 Db 202 -----DNNNIEELPN-----DVQAGSPYIIDISTKYN 231
 QY 412 SIHPEASTLSLVKLDVLTQNLTPPLAGLGLMLKLGKLNLSOAFSKSPFLRLIL 471
 Db 232 S-----LPNGLDENKKLRASVYRLKPLMLDFVILME 267
 QY 472 EPPYAYOCCEPYGMCASFPGKSGOWEADLH-----LDDEE 506
 Db 268 SLTFPSHCCAF-----ANLKRQISELHPICNKSILRODIDMTQIGQRYSLIDDEP 319
 QY 507 SSKRPLGLIAQANHNHYDQDELQLEMEDESKPSPVQSCPTPQPFKCEYLFESWGIRL 566
 Db 320 S-----YCGSDMMYNEFDYLCN-----EYVD-----VTCSPKPDAPFPCEDINGYNILRV 366
 QY 567 AVMAIVLISVLCNGLVLTVPAGPAPLPVYKVGAGIAGANTLTGISCGLASVALTF 626
 Db 367 LWFISILAITGNTTVLV-VLTTSQYKLTVPRLMCLNLAFAADLCIGYLLILIASVDHTK 425
 QY 627 GQFSEYGARWETGICGATGFLAVLGSPASVLLTLIAVQ-----GSVSCYRAVGS 680
 Db 426 SQYHNVAIDMOTGACCDAGFEYVASELSVYTLATLTLEWHITTHAMQLE-----KV 480
 QY 681 PELGVRVAGVGCIALAGIAALPLASVGEYASGLCLPVAPPGOPALGFTVALVMAN 740
 Db 481 QLRHIAVAVLG-WTFAPRAALPFIPIGSSYMKVSICLPM--DIDSPLSQLYVALVILN 537
 QY 741 SPCFLVAVAGATYIKLYCDLPGRGF-BAVWDCAMVRVAVMLIADGLLYCPVAFSPASMLG 799
 Db 538 VLAFFVIGCGYHIVLTVPNTIVSSSPYKIAKMAVLIITFDLCMAPISFPAISLAK 597
 QY 800 LRPVYPEAKSVLLVLPACLNPLVLTLPNPRDILRLRPAGDSGLAVAAAGEL 859
 Db 598 VPLITVSKAKILLVLEFPLNSCANFLVAIFTRNFRDRDFIL-----L 640
 QY 860 EKSSCDSTQALVAFSDVLDILEASEGRPPGLETYGFPVSYLLISCOQPGAPRLBSHCVE 919
 Db 641 SKFGCYEMQAQIYRT-----ETSSA-----THNF-HARKSHCS--SAPRTNSYLV 684
 QY 920 PEGNH 924
 Db 685 PL-NH 688

RESULT 8

QY 115 SIKIIMLONNCGIPAEALMELP-SLOSLRLANMLISLVERSEGLSSLRHMLDNA 173
 Db 26 SNRVLPCDSKYETLP-----DLPRNALELRFVLFKRLKRVIPKSPAGCGDLKRIISOND 81
 QY 174 LFE-IPVRLANMLPALQAMTL-ALNRISITPIYAPONLTSVVLVLLHNRRIOHGTSHFE 231
 A:Title: Folliotropin receptor precursor - human
 N:Alternate names: follicle stimulating hormone receptor (FSHR)
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence_revision 06-Sep-1996 #text_change 22-Jun-1999
 C:Accession: I57661; I56448; P01147; S30560; I57672; J01022
 R:Gromoll, J.; Dankbar, B.; Gudermann, T.
 Mol. Cell. Endocrinol. 102, 93-102, 1994
 A:Title: Characterization of the 5' flanking region of the human follicle-stimulating
 A:Reference number: I57661; M01D:95011044
 A:Accession: I57661
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-51 <GRO>
 A:Cross-references: GB:S73199; M1D:9685036; P1D:AA832071.1; P1D:9685037
 R:Gromoll, J.; Ried, T.; Holtgreve-Grez, H.; Nieschlag, E.; Gudermann, T.

C:Accession: JN0898; S36452
 R:Gomoll, J.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.
 Biochem. Biophys. Res. Commun. 196, 1066-1072, 1993
 A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of the
 A:Reference number: JN0898; M01D:94071854
 A:Accession: JN0898
 A:Molecule type: mRNA
 A:Residues: 1-695 <GRO>
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CA52463.1; PID:9396802
 A:Note: the authors translated the codon AGT for residue 488 as Arg
 C:Function:
 A:Description: receptor that mediates the biochemical effects of follicleotropin
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein;
 F:1-17/Domain: signal sequence #status predicted <Sig>
 F:18-695/Product: follicleotropin receptor #status predicted <PRH>
 F:71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
 F:194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
 F:167-387/Domain: transmembrane #status predicted <TM1>
 F:399-421/Domain: transmembrane #status predicted <TM2>
 F:444-465/Domain: transmembrane #status predicted <TM3>
 F:486-508/Domain: transmembrane #status predicted <TM4>
 F:529-550/Domain: transmembrane #status predicted <TM5>
 F:574-597/Domain: transmembrane #status predicted <TM6>
 F:609-630/Domain: transmembrane #status predicted <TM7>
 F:191-199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:556/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 10.2%; Score 515; DB 1; Length 695;
 Best Local Similarity 22.7%; Pred. No. 5, 8e-28;
 Matches 195; Conservative 117; Mismatches 298; Indels 248; Gaps 26;

```

QY 39 CHCEDGIMSDCESELGSAVPGDLPITAYLISMNLTLEQGFHHLFLELRIS 98
DB 23 CHCSNRFL-----QESKVEIRPSDL-PRNAI-----ELRV 54
QY 99 GNHLSHIPGAFSGLYSL-KILMIONNOGIGTAPBALMELPSQSRLDA--NLISLVE 155
DB 55 HTKLKVIQKGAISGFGLEKIEISQNDVLEIVDFVSNLPKHLRIEKANLTLINDE 114
QY 156 RSFELSRLHMLDNLATFIPRALNLTALINRISHPDAFQNLSTLYVL 215
DB 115 -----AFQNLPLNRYLLISMTGKHLDP----- 137
QY 216 HLHNNRIQHLGTHSEGLHNETLDLNTYKIQEPFVAIRLGLQELGFHNNIKAIPEK 275
DB 138 -VHK-----IHSFOKY-----LDDIQ-----DNINHTERN 163
QY 276 AWMGNPLQTHFYDNPLOFVRSAPQYLPKHLTSLGANDIOEPDLKGTSLLETLT 335
DB 164 SFVGLS-----FESVILMLKNIGIQ-----ELHNCAFNG-----TQDELNTL 200
QY 336 TRAGIRLPSGMCQQLPRLVLELSHNOIEELPSLHRQKLEIGLQNRIMEIGADTFS 395
DB 201 S-----DNNNLEELPN-----DVFH 215
QY 396 QUSLQALDLSMNAIRSIHPEAFSTLSLVKLDLTDNQLTLPPLAGLGLMLHLKGLMA 455
DB 216 GASGPVILDISKRIHS-----PSSYLENKLKLRASTRYN 251
QY 456 LSAFSDSPFKRLILEVPAVYACCPYGCASFKAAGWEAEDLH-----DD 504
DB 252 LKRLPSLEKLVALMEASLTPSHCAV-----ANMRQSLSELHPIKCNKSLROEVDY 303
QY 505 EESSRRRLGLLARQAEHHYDQDDELQLEMEDSKRPH--SVQSPTPPGPKCEVLEFEM 562
DB 304 MTQTRGQRSSLAEDNESSYSGRDFMTYAEFDYDLQNEVVDVTCSPKPDAPNCEIDILGYN 363

```

```

QY 563 GIRLAWAIVLVLGNGVLTLEFAGGPAPLPVYKVVVAGIAGANTLTGISCGLASVD 622
DB 364 ILKVLWIFISILATITGNITVLVLT--TTSQYKLTVPRLMNCNLAFAADLCIGYLLIASVD 422
QY 623 ALTFGESEKAWMERGLGCRATGFLAVIGSEASVLLTLAAVO-----CSVS 670
DB 423 IHRKSQHNVAIDMOTGAGDAGFFTFVASELSVYTLATITLERNHTTANMOLDCVH 482
QY 671 VSCVRAVYKSPISGSRVAGVIGLALAGLAALPLASVGEASPLCLPYAPPEGOPAL 730
DB 483 VR-----HAASVWGMWT--FAFAALFPIFGISSYMKVSIPLM--DIDSPLSQ 528
QY 731 GFTVALVMNSFCEFLVAGAYIKYCDLPGRDF--EAWDCAMVRHAMLIFADGILYCPV 789
DB 529 LYMSLVLTAVLAVVIGCGTHTYLVFRNPNTVSSSDTRIKRKNAMLTFDFLCMAPI 588
QY 790 AFLSPASMLGFPVTPPEAVSVLLVLPACINPLLYLFNFHFDLDRLRAPRAGDSG 849
DB 589 SFPAISALSKVPLITVSRAKILLVFLYPINSCANPFLYALFTNFRDPIILSKFG-- 645
QY 850 PLAVAAAGELKSSCDST 867
DB 646 --CYEKQAOITRTTSST 661

```

RESULT 10

Thyrotropin receptor precursor - dog
 N:Alternative names: thyroid-stimulating hormone receptor; TSH receptor
 C:Species: Canis lupus familiaris (dog)
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 13-Aug-1999
 C:Accession: A40077; S06933
 R:Parmentier, M.; Libert, F.; Meenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; Van S
 Science 246, 1620-1622, 1989
 A:Title: Molecular cloning of the thyrotropin receptor.
 A:Reference number: A40077; M01D:90084524
 A:Accession: A40077
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PA2>
 A:Cross-references: GB:M29957, NID:q164098, PIDN:AAA30901.1; PID:q164099
 R:Parmentier, M.; Libert, F.; Meenhaut, C.; Lefort, A.; Gerard, C.; Perret, J.; van S
 Nucleic Acids Res. 17, 10493, 1989
 A:Title: Nucleotide sequence of the dog thyrotropin receptor cDNA.
 A:Reference number: S06933; M01D:90098886
 A:Accession: S06933
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-764 <PA2>
 A:Cross-references: EMBL:X17146; NID:98449; PIDN:CA53026.1; PID:9850
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <Sig>
 F:21-764/Product: thyrotropin receptor #status predicted <MAT>
 F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F:201-225/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

Query Match 9.7%; Score 490.5; DB 2; Length 764;
 Best Local Similarity 22.7%; Pred. No. 3, 4e-26;
 Matches 211; Conservative 131; Mismatches 353; Indels 233; Gaps 30;

```

QY 4 PPGRLMILCALCASRRAGAPQPGPTACPAQC-QEGLMLSADCELSGSAVPG 62
DB 3 PPHLHALLALPRLSISGKGCPS-----PCEHQEDDFRT--CKDI----- 44
QY 63 DLDELAVLDSMNLTLEQGLFHLRLFLELRISGNHLSHIPGAFSGLYSLKILMIO 122
DB 45 -----HRIPTLPST-QTLKFL-----TQKLPISAFENLPIISRIVLS 84

```

```

QY 123 NNQIGIPAEALMELPSLQSLRLDANLISIVPERSEEGSLRIHML-DNALTEIPYRA 181
Db 85 -----IDATLQRL-ESHSEFWLSKMTHEIRNTSLSDPDA 121
QY 182 LNNPALQAMTLALNRISHIDY-----ARQNL-TSLVY 214
Db 122 LKELPLKFLGIDFTMTGCVFDPVAVKYSTDVFEILEITDNPYASIPANAFQGLCNETLT 181
QY 215 LHLNNRIQHLGTHSEFEGHNLFTLDNLYNKLOEFPVAIRTLGLQELGPHNNNIKAPE 274
Db 182 LKLTNNGFTSIQGHAFNGT-KLDVYLNKNKY-----LSAIDK 218
QY 275 KAFMGNPLQTHIFYNDP-----IOFVGRSAF-QYLPKLTLSLNGAMDOIEFPDLKGT 328
Db 219 DAGG-----VYSGPTLADVSTSVTALPSKGLHKLKELIARNTWLKLP-----L 265
QY 329 SLEILTLFRAGIRLPSGMCQQLPRLVLELSNQIEELPSLHRCQLEIGLOHNRIVE 388
Db 266 SLSEFLHLTRADLS-YPSHCC-----AFKNOKKIRGI----- 295
QY 389 IGADTFQSLSLQALDLSMNAIRSHIP-EAFSTLSLVKLDITDNLQTLPLAGLGLMH 447
Db 296 -----LESIMCNESISIRSFQRKSVNLTNGPFDQYEYEL-----GDSH 334
QY 448 LKLGKGLALSOAFSKDSFPEKRLILEVPYAYOCCPYGMCASFPGASGOWEADLHDDDES 507
Db 335 AGYKDN---SOPQDDSNHYV-----FPE-----EOEDEILFGQE 369
QY 508 SKRPLGLARQENHYDODLELOEMEDSKRHPSPVOCSPTPGPKPCEYLFESGIRLA 567
Db 370 LKNPEETLQAFDSHYDYV-----CGNEDMCTPRSDENPCEDIMGYKFLKIV 420
QY 568 VWAILVLSLNGVLLVTFVAGGAPAPLPVKEVVGATAGANTLTGISGLASVALTFG 627
Db 421 VMFVSLALGNVFLVILYLSHKLTVP-RFLMCNLAFADFCGMYLLTLASVDLYTHS 479
QY 628 QFSEKGAWEGLGCRATGFLAVLGSSEAVLLTLTAOVCSVSVCRAVGSLSGSVR 687
Db 480 EYVNHAIWMQPGCNTAGFEVTFVASELSYTLVITLERMYATTFMRDLKRLNHAY 539
QY 688 AGVLGCLALAGLAAALPLASVGEYSGLPLCLPAPRPGOPALGFTVALVMNSFEFLVY 747
Db 540 AIMGWVCCFELALPLPVGISYAKVSTCLPM-DTEPLALAYIIVLLINVAFLIV 597
QY 748 AGVYIKLYCPLRGDFE-AVWDCAMVRHVAWLTFADGILYCPVAPISFASMLGLFPVTEP 806
Db 598 CSCYVYKITYTVRNPQYNPGDKDKRIAKRMAVLFTDFMCMAPISFALSLMKNKPLITVY 657
QY 807 AVKSVLLVLPPLACNPLLYLYLEFNHFRDRLRPRAGDSGLAYVAAAG----- 857
Db 658 NSKILVLYPLNSCANPLPLYALFTKAFQRDVFLLSKFGICKRQAOAVRGQVSPKNSA 717
QY 858 --ELENSCDSSTQALVAFSDVDLILEAS 883
Db 718 G10IQKVTDRMROSLPNNODEYELLENS 745

```

RESULT 11

148882

thyrotropin receptor precursor - mouse

N:Alternate names: thyroid-stimulating hormone receptor; TSH receptor

C:Species: Mus musculus (house mouse)

C:Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 13-Aug-1999

R:Stein, S.A.; Oates, E.L.; Hall, C.R.; Grumbles, R.M.; Fernandez, L.M.; Taylor, N.A.; F

Mol. Endocrinol. 8, 129-138, 1994

A:Title: Identification of a point mutation in the thyrotropin receptor of the hyl/hyt H

A:Reference number: A54271; MUID:94224232

A:Accession: 148882

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-764 <RES>

A:Cross-references: EMBL:U02602; NID:g575923; PIDN:AAB60455.1; PID:g575924

C:Genetics:

A:Gene: TSHR

C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat

C:Keywords: G protein-coupled receptor; transmembrane protein

F:53-76/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>

F:77-101/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

F:102-126/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>

F:127-151/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>

F:152-176/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>

F:179-200/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

F:201-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

Query Match 9.6%; Score 486; DB 2; Length 764;

Best Local Similarity 21.7%; Pred. No. 7.1e-26;

Matches 211; Conservative 132; Mismatches 346; Indels 284; Gaps 30;

QY 4 PGLRALMCAALCASRAGAPQPGPTACPAQHC-OEDGIMLSNDSCELSAVPG 62

Db 3 PGLLILVLLALSRSLR-----CKEASPCCECHQDDFRVT--CKEL----- 44

QY 63 DLDPFLAVLDLSMNNITPELOGLFHHLRLEELRLSGNHSHTPGAFSGLYSLKIMLQ 122

Db 45 -----HRPSLPS-----TQTLKLTFTHTKTPISLAFSSLSRPNISRIYLS 84

QY 123 NNQIGIPAEALMELPSLQSLRLDANLISIVPERSEEGSLRIHML-WIDNA 173

Db 85 -----IDATLQRLP-HSFYNLSKMTHEIRNTSLTYIDPA 121

QY 174 LTEIP-----VRAINPALQAMTLAL-----NRISITPDYAFQNL-TSLVY 214

Db 122 LTELPLKFLGIDFTMTGIRFPDLTKIYSTDIFILEITDNPYMTSVENAFQGLCNETLT 181

QY 215 LHLNNRIQHLGTHSEFEGHNLFTLDNLYNKLOEFPVAIRTLGRLQELGPHNNNIKAPE 274

Db 182 LKLTNNGFTSIQGHAFNGT-KLDVYLNKNKY-----LTAIDN 218

QY 275 KAFMGNPLQTHIFYNDP--IOFVGRSAFOYLPK-----LHTLSLNGAMDOIEFPDLKGT 328

Db 219 DAGG-----VYSGPTLADVSSVTVATLPSKGLHKLKELIANDWTMLKLP-----L 265

QY 329 SLEILTLFRAGIRLPSGMCQQLPRLVLELSNQIEELPSLHRCQLEIGLOHNRIVE 388

Db 266 SLSEFLHLTRADLS-YPSHCCAFKNOKKIRGILIESLMCNESIRILRQKSVNIIIRGPIYQ 324

QY 389 -----IGADTFQSLSLQALDLSMNAIRSHIPAFSTLSLVKLDITDNLQTLPL 438

Db 325 EYEDPQGNVSG--YKQNSKFQ-----ESPNSHNYV----- 354

QY 439 LAGLGLMHLKGLMALSQAFSKDSFPEKRLILEVPYAYOCCPYGMCASFPGASGOWEAE 498

Db 355 -----FPE-----EOE 360

QY 499 DLHDDDESSKRPGLGLARQENHYDODLELOEMEDSKRHPSPVOCSPTPGPKPCEYL 538

Db 361 DEAVGFOELKNPQDETLQAFESHYDYV-----CGDNEDMCTPRSDENPCEDL 411

QY 559 FESWGIRLAWAILVLSLNGVLLVTFVAGGAPAPLPVKEVVGATAGANTLTGISGLY 618

Db 412 MGYEFLRIVVWFVSLALGNITVLLTLSHKLTVP-RFLMCNLAFADFCGVYLLLI 470

QY 619 ASVDALTFGQFSEYGAWEGLGCRATGFLAVLGSSEAVLLTLTAOVCSVSVCRAAG 678

Db 471 ASVDLTTHSEYVNHAIWMQTPGCGNTAGFEVTFVASELSYTLVITLERMYAITTFMRDL 530

QY 679 KSPSLSVRAGVGLGCLALAGLAAALPLASVGEYSGLPLCLPAPRPGOPALGFTVALVM 738

Db 531 RKRLRHATTIAGGAVSCFPLALPVGISSYAKVSTCLPM-DTEPLALAYIVLVLL 588

QY 739 MNSFCFLVVAAYIKLYCDL-----PRGDEFAVWDCAMVRHVAWLTFADGILYCPVAPL 792

Db 589 LNVVAFVVVVCSCYKITYTVRNPQYNPRDK-----DTIARMAVLFTDFMCMAPISFY 643

A:Title: Lutropin-choriogonadotropin receptor: an unusual member of the G protein-coupled
 A:Reference number: A41343; MUID:89332512
 A:Accession: A41343
 A:Molecule type: mRNA
 A:Residues: 1-700 <MCF>
 A:Cross-references: GB:M26199; NID:g205178; PIDN:AAA41528.1; PID:g205179
 R:Putnam, M.L.; Minegishi, T.; Buczeko, E.S.; Delgado, C.J.; Zhang, R.
 J. Steroid Biochem. 33, 715-720, 1989
 A:Title: Characterization and structure of ovarian and testicular LH/hCG receptors.
 A:Reference number: A61453; MUID:90097014
 A:Accession: A61453
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 27-33, 'X', 35-37, 'X', 39, 'X', 41-44 <DUF>
 R:Roche, P.C.; Ryan, R.J.
 J. Biol. Chem. 264, 4636-4641, 1989
 A:Title: Purification, characterization, and amino-terminal sequence of rat ovarian recep
 A:Reference number: A32460; MUID:89174723
 A:Accession: A32460
 A:Molecule type: protein
 A:Residues: 27-32, 'LX', 35-37 <ROC>
 C:Genetics:
 A:Insertions: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
 C:Superfamily: glycoprotein hormone receptor: leucine-rich alpha-2-glycoprotein repeat h
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
 F:54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:368-389/Domain: transmembrane #status predicted <TM1>
 F:400-422/Domain: transmembrane #status predicted <TM2>
 F:444-466/Domain: transmembrane #status predicted <TM3>
 F:489-511/Domain: transmembrane #status predicted <TM4>
 F:530-551/Domain: transmembrane #status predicted <TM5>
 F:575-598/Domain: transmembrane #status predicted <TM6>
 F:610-631/Domain: transmembrane #status predicted <TM7>
 F:103-178, 199, 295, 303, 317/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:647, 648/Binding site: palmitate (Cys) (covalent) #status predicted
 F:661/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 9.4%; Score 474; DB 2; Length 700;
 Best Local Similarity 21.5%; Pred. No. 4,3e-25;
 Matches 203; Conservative 130; Mismatches 322; Indels 288; Gaps 30;

QY 5 PGLRALMLCAL-----CASRRAGAPQPGPGPTACAPACCHQEDGIMLSADCSSEGLS 58
 DB 6 PALROLVLAVLLKPSQLSRELSG-----SRCPECDACPDGAL----- 46
 QY 59 AVPGDDLPFLAYDLDSNNNLTELQGLFHHIRFLLERLSGNHLSHIPGQAFSGLYS-LK 117
 DB 47 RCPG---PRAGLARLSLTYLP-----VKVIPSQAFRGILNEVYK 81
 QY 118 ILMQNNQGGIPAEALMELPSLOSLRD--ANTLSVPERSEGLSLRHLMDNALT 175
 DB 82 IEISQSDSLERIEANNDNLINSELILQNTKNLIYEP-GAFTNLPRLKTLISICNTGIR 140
 QY 176 EIP-VRALNNLPALQAMTALN-RISHIPDYAFQNLTS-LVYLHLHNNRIQHLGTSFEG 232
 DB 141 TLDPVTKISSSEFNLEICDNLHITIPGAFQGMNNESTLKYNGFEEOVSHAENG 200
 QY 233 LHNLETDLNKNKIQEPPVAIRTLGRLOELGFHNNNIKAIPKAPMGNPDLQTIHFIDNP 292
 DB 201 T-TLISLELEN----- 211
 QY 293 IOEFGRAFOYLPKLTHTLSINGAMDIQEPDCKTSTLSLEITLRAGIRLPSGMCOQLP 352
 DB 212 -----TYLEKMHSGAFQATG-----PS----- 229
 QY 353 RLAVLELSHNOIELPSLHRCQKLEIEIGLOHNRIMEIGADTFQSLSLQALDLSWNAIRS 412

DB 230 ---ILDISSTKLQALPS-----HG-----LESIQTL----- 252
 QY 413 IHPEAFSTLHSLVLDLIDNQLTTLPLAGLOGIMHLKGNLALSQANSKSPFKRLTLE 472
 DB 253 -----TALSYSILKTLP-----SKREFTSLVAT 276
 QY 473 VPYAVQCCPYGMCASFQKASQWEAEDHLDESSKRPLGLLAQAEHN--YQDDLEL 530
 DB 277 LTYSHCCAFNNLR-----KEQNSFSFIENFSQCESTYKKNANELLYSAIFERN 328
 QY 531 QLEMED-----SKHPVSQCSPTPGEPKPCYLEFESWGIRLAWAIVLLSVLCNGVLIT 585
 DB 329 ELSGWDYDYGCSF-KTLOCAPEPDAPFPCEDINDGAYLARVLIIMINILATFGN-LTVLF 386
 QY 586 VFAGGAPLPPVKRVVGAIAQANLITGSCILLASVDLITGQSEYCARRETTGCGCAT 645
 DB 387 VLTFSRKLTVPFRFLMCMLSFADFCMGLYLLILASVDSQTYGNHAIMDQTSGCCAA 446
 QY 646 GFLAVLGSSEASVLLDITLAAVQCSVSCVRAYGKSPISGVRAGVIGCLALAGLAAALPL 705
 DB 447 GFTYFASELSTVLTITVTLERWHITTYAVQDDKRLRHAIPIMLGWLESTLITATMPL 506
 QY 706 ASVGEYGASPLCLPYAPPEGQPALGFTVALVMNSFCFLVAVAGAYIKLYCDLPRGDEA 765
 DB 507 VGISNMYKVSICLPM--DVESTLSQVYLLITLILNVAFAVVIACACYIRIYFAVQNPBELTA 564
 QY 766 V-WDCAMVRHVAWLIFADGLIKCPVAFISFASMLGFVTPENKVSULLVPLPACINP 824
 DB 565 PNKDKTKAKKMAILIFDFTCMAPISAEFAISAERKVPILITVNSKILVLVYVNSCANP 624
 QY 825 LLYLLFNFHFDLRLRPRAGDSGLPAAVAAAGLELSSCDSTQALVAFSDVDLILEASE 884
 DB 625 FLVAIFTFAPRDFLLLSRFG-----CKRRA-----E 653
 QY 885 AGRPPELTY-----GFP-----SVTLISQQPGAPR 911
 DB 654 LYRRKEFSAYTSNCKNGPFGASKPSQATLKLSTVHCQCPIDPR 696

Search completed: October 23, 2002, 10:41:36
 Job time : 35 secs

v
.
.
.
.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52; Search time 15 Seconds
(without alignments)
2496.120 Million cell updates/sec

Title: US-09-851-595-11

Perfect score: 5068
Sequence: 1 MPEPGLRLMLCALCASR.....GGLSGGCGFPQSLAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.5	10.5	692	1	FSHR_MOUSE
2	526.5	10.4	695	1	FSHR_BOVIN
3	525	10.4	694	1	FSHR_HORSE
4	524.5	10.3	695	1	FSHR_SHEEP
5	523	10.3	695	1	FSHR_PIG
6	518.5	10.2	692	1	FSHR_RAT
7	518	10.2	695	1	FSHR_HUMAN
8	515	10.2	695	1	FSHR_MACACA
9	504.5	10.0	687	1	FSHR_EQUUS
10	490.5	9.7	764	1	FSHR_CANPA
11	486	9.6	764	1	FSHR_MOUSE
12	479	9.5	763	1	FSHR_BOVIN
13	474	9.4	760	1	FSHR_RAT
14	473.5	9.3	764	1	FSHR_SHEEP
15	472	9.3	700	1	FSHR_MOUSE
16	469	9.3	701	1	FSHR_BOVIN
17	467	9.2	696	1	FSHR_PIG
18	466	9.2	693	1	FSHR_CHICK
19	464.5	9.2	764	1	FSHR_HUMAN
20	462.5	9.1	764	1	FSHR_RAT
21	455.5	9.0	603	1	ALS_RAT
22	444.5	8.8	676	1	LSHR_CALJJA
23	439.5	8.7	699	1	LSHR_HUMAN
24	439	8.7	605	1	ALS_PAPHA
25	438.5	8.7	603	1	ALS_MOUSE
26	437.5	8.6	605	1	ALS_HUMAN
27	422.5	8.3	567	1	GPV_RAT
28	411	8.1	925	1	GLHR_ANGEL
29	397	7.8	1480	1	SLIT_DROME
30	394	7.8	538	1	LSHR_SHEEP
31	375	7.4	560	1	GPV_HUMAN
32	368	7.3	567	1	GPV_MOUSE
33	364.5	7.2	1315	1	CHAO_DROME

34	363	7.2	536	1	CBP8_HUMAN	P22792 homo sapien
35	363	7.2	366	1	GARP_HUMAN	Q14392 homo sapien
36	357	7.0	662	1	LSHR_CHICK	Q90674 gallus gall
37	349	6.9	713	1	GACL_HUMAN	Q75325 homo sapien
38	348.5	6.9	1115	1	GPCR_TLMST	P46023 lymnaea sta
39	348	6.9	782	1	CHAO_TRICA	P82963 trilobium c
40	346	6.8	1049	1	TLR7_HUMAN	Q99414 homo sapien
41	340.5	6.7	1039	1	YR71_CAEEL	Q09564 caenorhabdi
42	337.5	6.7	582	1	SHO2_HUMAN	Q9ug13 homo sapien
43	326.5	6.4	1050	1	TLR7_MOUSE	P58681 mus musculu
44	326	6.4	582	1	SHO2_MOUSE	O68520 mus musculu
45	324.5	6.4	905	1	TLR3_MOUSE	Q99b01 mus musculu

ALIGNMENTS

RESULT 1
FSHR_MOUSE STANDARD; PRT; 692 AA.
AC P35378: O9QW8: Q9DAC2:
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictropin receptor).
GN FSHR
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Testis;
RA Tena-Sempere M., Manna P.R., Huhtaniemi I.T.;
RT "Molecular cloning of the mouse follicle stimulating hormone receptor complementary deoxyribonucleic acid: functional expression of alternatively spliced variants and receptor inactivation by a C566T transition in exon 7 of the coding sequence."
RT submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I., Saito T., Okazaki Y., Gojopori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schiml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P., Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
[3]
RN SEQUENCE OF 1-51 FROM N.A.
RP MEDLINE=93093308; PubMed=1459341;
RA Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Springer R.;
RT "The murine luteinizing hormone and follicle-stimulating hormone receptor genes: transcription initiation sites, putative promoter sequences and promoter activity."
RL Mol. Cell. Endocrinol. 88:55-66(1992).

CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/SH/SH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF095642; AAC67559.1; -
 DR EMBL: AK016635; BAB30351.1; -
 DR EMBL: S49632; AAB24401.1; -
 DR EMBL: M87570; AAB37641.1; -
 DR GCDDB: GCR_0304; -
 DR MGD: MGI:95583; Fshr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 DR Phosphorylation; Repeat; Leucine-rich repeat.
 KW SIGNAL 1 17
 FT CHAIN 1 692
 FT DOMAIN 18 365
 FT TRANSMEM 366 385
 FT DOMAIN 387 397
 FT TRANSMEM 398 420
 FT DOMAIN 421 442
 FT TRANSMEM 443 464
 FT DOMAIN 465 484
 FT TRANSMEM 485 507
 FT DOMAIN 508 527
 FT TRANSMEM 528 549
 FT DOMAIN 550 572
 FT TRANSMEM 573 596
 FT DOMAIN 597 607
 FT TRANSMEM 608 629
 FT DOMAIN 630 692
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 441 516
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT CONFLICT 436 436
 SQ SEQUENCE 692 AA; 77769 MW; 4B5722918053A44 CRC64;
 Query Match 10.5%; Score 532.5; DB 1; Length 692;
 Best local Similarity 23.7%; Pred. No. 4.1e-29;
 Matches 202; Conservative 124; Mismatches 293; Indels 235; Gaps 30;

DB 82 VLEVEAEVFNLPRLHRIERIKANNLTYNEAONLPSLYLSTNGIKHL-----P 136
 QY 232 GLHNLLETDLNKNKLOEPPVAIRTLGLOEIGHNNNIKAIPKAFMGNPLDTTHFYDN 291
 DB 137 AFHKIOSLQ-----KVLIDIOD--NINIIHARSFMIS--FES 172
 QY 292 PLOFVGSAFQYLPRLHTLSNGAMDIOEPDLKDTLSLELTTLRAGIRLLPSGCGQL 351
 DB 173 VILNKNKGIO---EINHCARNG-----TOLDENLS----- 201
 QY 352 PRLVLELSHNOIELPESLHRCQKLEIGLOHNRITWEGADTFSSLOALDSMNAIR 411
 DB 202 -----DNNLEBLDP-----DVFQASGVVVIDISRTKY 231
 QY 412 SIHPEASTLSLVKLDLTDNQLTLPAGIGGLHLKGLMALSOAFKSDSPKRLTL 471
 DB 232 S-----LPNHGLENLKLLRARSYYRLKLPDLKFYMLIEA 267
 QY 472 EVPYAYOCPEYGMCAFFKASGOWEAEIDLDDDESSKRPGLGLARQAEHNDODDEL- 530
 DB 268 SLITPSHCAP-----ANWRKOTSELHPTCNKSISR-----QDIDMT 305
 QY 531 -----QLEMEDSKP-----HPSVQCSPTPGPFKPCXYLFESMGIR 565
 DB 306 OPQDQVSLVDDEPSYSGSDMLYSEFDYDLCNEFDVDTGSPKPAFNPCEIDIMKYNILR 365
 QY 566 LAWAIVLISLVGLGLVLLTYFAGGPAPLPYKVVYCAIACANLITGISGLIASVDALT 625
 DB 366 VLIWFISILAITGNTVLY-VLTTSQYKLVPRFLMCLAPADICIGYLLILIASVDIHT 424
 QY 626 FGFSESGAWMETGLGRANGFLAVLSEASVLLITLAAY-----QCSYSV 671
 DB 425 KSOYHNVAIDMQGAGGADGFFTFVASELSV--YTILAITLERHNTTHAMOLECKYOL 482
 QY 672 SCYRANGKSPSLGSVRAGVGLCAGLAAALPLASVGEYASPCIGPYAPPEGOPALG 731
 DB 483 -CHAA-----SIM--VIG-WAFAPAAALFPIFGISSYKAVSICLPM--DIDPLSQL 528
 QY 732 FYVALVMNSECFFVYVAVIYIKYCDLPRGDF-EAVVMDCAVRVAVAMLIFADGLLYCYVA 790
 DB 529 YVVALVLYNALVAVYIGCTHTIYLVNRPVYSSRRTKAKMAWLTIFDFLCMAITL 588
 QY 791 FLSPASMLGFPYTPPEAVKSVLLVPLPACLPNPLLYLLEPHFRDILRLRPAGDSGP 850
 DB 589 FFAISASLKVPLTIVSKAKITLVLYFPINSCANPFLVAIFTKNRRDPEVL----- 639
 QY 851 LAFAAGLEKSSCDSTQALVAFSDVLLLEASAGRPGLIETYGFPSSVLLISCOQPGAP 910
 DB 640 -----MSKFGCYEVOAQIYKTEISSITHNFSRRNP-----CS--SAP 675
 QY 911 RLEGSHCVPEEGNH 924
 DB 676 RVYNSYVLPVPL-NH 688
 RESULT 2
 FSHR_BOVIN
 ID FSHR_BOVIN STANDARD; PRT; 695 AA.
 AC P35376;
 DT 01-JUN-1994 (rel. 29, Created)
 DT 01-JUN-1994 (rel. 29, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (follicleotropin
 DE receptor).
 GN FSHR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.

CC	STRAIN-HOLSTEIN; TISSUE-Ovary; and Testis;
RX	MEDLINE=9512719; PubMed=7826612;
RA	Houde A., Lambert A., Saumande J., Silversides D.W., Justier J.G.;
RT	"Structure of the bovine follicle-stimulating hormone receptor
RL	molecular cloning, expression and characterization in ovine tissues";
CC	1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC	FSH/LSH/TSH SUBFAMILY.
CC	-1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/or send an email to license@isb-sib.ch).
CC	EMBL; L22319; AAC37324.1; -.
DR	HSSP; P23945; IXUN.
DR	GCRDB; GCR_0766; -.
DR	InterPro: IPR000276; GPCR_Rhodpsn.
DR	InterPro: IPR001611; LRR.
DR	InterPro: IPR000372; LRR_Nterm.
DR	Pfam; PF00001; 7tm1.1; 1.
DR	Pfam; PF00560; LRR; 5.
DR	Pfam; PF01462; LRRNT; 1.
DR	PRINTS; PR00373; GLYCHROMONER.
DR	PRINTS; PR01143; FSHRECEPTOR.
DR	SMART; SMD0013; LRRNT; 1.
DR	PROSITE; PS00237; G_PROTEIN_RECP_F1.1; 1.
DR	PROSITE; PS50262; G_PROTEIN_RECP_F1.2; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.
FT	SIGNAL 1 17 POTENTIAL.
FT	CHAIN 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
FT	DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 367 387 1 (POTENTIAL).
FT	DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 399 421 2 (POTENTIAL).
FT	DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 444 465 3 (POTENTIAL).
FT	DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 486 508 4 (POTENTIAL).
FT	DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 529 550 5 (POTENTIAL).
FT	DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 574 597 6 (POTENTIAL).
FT	DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 609 630 7 (POTENTIAL).
FT	DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
FT	REPEAT 44 68 LRR 1.
FT	REPEAT 69 93 LRR 2.
FT	REPEAT 119 143 LRR 3.
FT	REPEAT 170 192 LRR 4.
FT	REPEAT 193 216 LRR 5.
FT	REPEAT 218 240 LRR 6.
FT	DISULFID 442 517 BY SIMILARITY.
FT	CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE 695 AA; 78084 MW; 18F9DFEFC046380D CRC64;
Query Match	10.4%; Score 526.5; DB 1; Length 695;
Best Local Similarity	26.1%; Pred. No. 1.le-28;
Matches 170; Conservative 108; Mismatches 290; Indels 83; Gaps 15	

Dd	50	ELRVLVTLKLVIPKGAASGSGDLEKIEISQNDVLEIVTEANVSNLPKLTHERIKERANNLL	109
Qy	320	EE-PDLKGTSTLELTLITLRAGCIRLPLSPGMCQOQLPRVLELISHNQEEELPSLRKOKLEE	378
Dd	110	YIDPD-----AFQNPINIRYLLISWTGKILPAVHKIQSLQK	146
Qy	379	I--GLOHN-RWIEIGADTFQSLS-SIQALDLSNNAIRSHIPEAF-----STL	421
Dd	147	VLLDIQDININIHIERKSPFAGLSPESTAYVLSKNGIOELHNCAFNQTQDLELNSQNSML	206
Qy	422	HSL-----VKIDFTDNDITTLPLAGLGLMHLKGNLALSQAFSKDSPFKLRI	470
Dd	207	EELPDVDFQGASGVVIDISTRIRSPYSGLENLKKLRKASTYRLKLPDLSLEKVTLYE	266
Qy	471	LEVPAVAVOCCPYGNGCAEFPAAGQWEMEDH-----LDDESSKRPLGLLARQA	519
Dd	267	ASLTPPHSCAF-----ANMRQTSIDLPLICNKSILRQEVDDMTQARGQRYSLAEDD	318
Qy	520	ENHYDQDLDLELEMEDESKPDP--SVQCSPTPGPFKCEYLFEESWIGIRLAWAVILSVL	577
Dd	319	EPYAKGPDVWYSEEDYDLCEVVDVDCSEPDPAFNPCEIDIMDDILIRYLWISILATIT	378
Qy	578	CNGVLTLVFPAGGAPRLPRYKFFVGALAGANTLTGISCGLLASVDALTGEOSEYCARBE	637
Dd	379	GNITLV--ILTTSQYKLTVPFRFLCNLAFADLCIGYLLTLASVDVHTKEYNVAIDMQ	437
Qy	638	TGLGCRATGFLVAVGSEASVLLTLTAAVQCSVSVCYRAVKGSPISGVNAGVGLA	697
Dd	438	TGAGCDAGGFTVVASASELSVTLTATILERNHTITHMQLECKVQLRHNASIMLVGMPA	497
Qy	698	GIAALPLASVGEYGASPLCLPYAPRPGQPAALGFTVAYLMANSFCPLVYAGIKLYCD	757
Dd	498	FAVALFPFIFGISTYMKVKSICLPM--DIDSPLSQLYWSSLVNLVLAWVCGCYTHIYLT	555
Qy	758	LPGRGF-EAWMDCAVVRVAMVLTIFADGLTLCPVAFILSFASMLGLFPTPEAVSVLVVL	816
Dd	556	VANPNITSSSDPTKAKKAMALITFDLCAPRISFPALSNLKVPLTYSKSLILVLFY	615
Qy	817	PLPACINPLLYLLFNPFRDDLRLRPRAGDSGLPAYAAGLEKSSCDST	867
Dd	616	PINSCANFLVAFITKFNRRDFILLSKFG----CYEVOAQYRSETST	661
RESULT 3			
FSHR_HORSE			
ID	FSHR_HORSE	STANDARD;	PRT; 694 AA.
AC	P47799;		
Dt	01-FEB-1996 (Rel. 33, Created)		
Dt	01-FEB-1996 (Rel. 33, Last sequence update)		
Dt	16-OCT-2001 (Rel. 40, Last annotation update)		
De	Follicle stimulating hormone receptor precursor (FSH-R) (folliotropin receptor).		
GN	FSHR.		
OS	Equus caballus (Horse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.		
OX	NCBI_TaxID=9796;		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=Testis;		
RX	MEDLINE=94256980; PubMed=8198575;		
RA	Robert P., Amselem S., Christophe S., Benifla J.L., Bellet D.,		
RT	Komen A., Bidart J.M.;		
	"Cloning and sequencing of the equine testicular follitropin		
RL	receptor."		
RL	Biochem. Biophys. Res. Commun. 201:201-207(1994).		
CC	-I- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY		
CC	OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE		
CC	ADENYLATE CYCLASE. AMONG ALL MAMMALIAN FSH RECEPTORS, ON THE HORSE		
CC	RECEPTOR DOES NOT BIND LH/CHORIONIC GONADOTROPHIN (CG).		
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
CC	FSH/LSH/TSH SUBFAMILY.		

291 NPIDPVGSRSAQYLPKHTLSLUGAMDQIEFPDLEKGTSTSEILTLTRAGIRLLPSPGMCQ 350

\mathbb{P}^1

[illegible]

SEQUENCE FROM N.A. (ISOFORM FSH-R2), AND CHARACTERIZATION.
 RE STRAIN-DORSET-LEICESTER-SUPFOLK 1; TISSUE-Testis;
 RC MEDLINE=98031015; PubMed=9364440;
 RA Varney T.A., Jiang L., Khan H., MacDonald E.A., Laird D.W.,
 SA Sairam M.R.;
 RT "Molecular cloning, structure, and expression of a testicular
 RT follicleotropin receptor with selective alteration in the carboxy terminus
 RT that affects signaling function.";
 RT Mol. Reprod. Dev. 48:458-470(1997).
 RL [4]
 RN SEQUENCE FROM N.A. (ISOFORM FSH-R3), AND CHARACTERIZATION.
 RP STRAIN-DORSET-LEICESTER-SUPFOLK; TISSUE-Ovary;
 RC MEDLINE=20391225; PubMed=10527886;
 RA Babu P.S., Jiang L., Sairam A.M., Touyz R.M., Sairam M.R.;
 RT "Structural features and expression of an alternatively spliced growth
 RT factor type I receptor for follicleotropin signaling in the developing
 RT ovary.";
 RT Mol. Cell Biol. Res. Commun. 2:21-27(1999).
 RL [5]
 RN SEQUENCE OF 1-51 FROM N.A.
 RP MEDLINE=98031017; PubMed=9364442;
 RA Sairam M.R., Subbarayan V.S.R.;
 RT "Characterization of the 5' flanking region and potential control
 RT elements of the ovine follicleotropin receptor gene.";
 RL Mol. Reprod. Dev. 48:480-487(1997).
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
 CC of isoform FSH-R1 is mediated by G proteins which activate
 CC adenylylate cyclase. Isoforms FSH-R2 and FSH-R3 also bind FSH, but
 CC this does not result in activation of adenylylate cyclase. Isoform
 CC FSH-R3 may be involved in calcium signaling. Plasma membrane
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
 CC (isoforms FSH-R1 and FSH-R2); Cell surface (isoform FSH-R3).
 CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; FSH-R1 (shown here), FSH-R2,
 CC FSH-R3 and FSH-R4; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Isoform FSH-R3 is expressed in ovary and
 CC testis, but not in kidney.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/TSR SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).

TRANSMEM 367 387
 FT DOMAIN 368 398
 FT TRANSMEM 399 421
 FT DOMAIN 422 443
 FT TRANSMEM 444 465
 FT DOMAIN 466 485
 FT TRANSMEM 486 508
 FT DOMAIN 509 528
 FT TRANSMEM 529 550
 FT DOMAIN 551 573
 FT TRANSMEM 574 597
 FT DOMAIN 598 608
 FT TRANSMEM 609 631
 FT DOMAIN 631 695
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 442 517
 FT CARBOHYD 191 199
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT VARSPPLIC 126 133
 FT VARSPPLIC 135 695
 FT VARSPPLIC 224 259
 FT VARSPPLIC 260 695
 FT VARSPPLIC 643 670
 FT VARSPPLIC 671 695
 FT SEQUENCE 695 AA; 78237 MM; BFB75D89D8C0D4B CRC64;
 SQ
 Query Match 10.38; Score 524.5; DB 1; Length 695;
 Best Local Similarity 26.48; Pred. No. 1; se-28;
 Matches 175; Conservative 101; Mismatches 273; Indels 113; Gaps 18;
 261 ELGFNNNNKRIAPKAPFNGNPLQTHFYDNP1-QFYGRSAFYQPKLHTLSLGANDIQ 319
 50 ELRFVLTIRVTPGAFSGFDLEKIELSQNDVLEIVANFSLPKLHEIRIKANNLL 109
 320 EF-PDLKGTSTLEILLTRAGIRLLPSGMCQQLPRLVLELSHQOIEPLSLHCQKLE 378
 110 YIDPD-----AFQNTPLRYLLISNGIKHLPVHRKIQSLQK 146
 379 I--GIQHN-RWIEIGADTFESOLS-SLOALDLSMNAIRSIHPEAF-----STL 421
 147 VLLDIQDININITHVERNSFMGISTFESMTWLSKNGIQELHNCATFNGTOLDLNLSDNSNL 206
 422 HSL-----VKLDLTQNLTLPLAGIGIMLKLKGNLALSOAFSKDSFPKIRI 470
 207 EELPNDVFGASGPVILDISRTIRSLPSYGENLKLIRAKSTYHLKLPLSEKRVTLVE 266
 471 LEVPAVOCPPYGMCAFPKASGQWEADLH-----LDDDESSKRPGLLARQA 519
 267 ASLTYSHCCAF-----ANWRQTSDLHPICNKSILRQEDVDMTQAKGORISLADD 318
 520 ENHYDDDELQLEWEDS--KPEYSVCSPTGPPEKPCSEYFESWGLRVAWVAIVLSVL 577
 319 EPSYAKGDMKSEFPYDLCSEVVDYTCSPEDAFNCPCDINGYDLRLWIFISILAIT 378
 578 CNGVLLTVFAGGPAPLPVKRYVGAIAAGANTLTIGISGLASVDAVLTGQFSEYGARE 637
 379 GNILVIV-ILITSQKLTVPREFLMCMIAFDALCTGITYLLILASVVDHRSQTHNAIDMQ 437
 638 TGAGCATATGLAVLGEASVLLTLAAVO-----CSVSVSC---VRAYGKSPSLGSVRA 688
 438 TGAGCDAGGFTVFASBLASYTTITATLERWHITIAMQLECKVHRHAASIMLVGV-- 495
 689 GVLGIALIAGIAAALPLASVGEASGPLCLVPAPBGOPAALGFTYALVMMNSFCLVYA 748

Db 207 EELPNDVFGASGPVILIDISRTIRHSLPSYGLENLKRLAKSTYNNLKKPLSEKFTVLTME 266
 QY 471 LEVPYAVQCCPYKMGASFFKASGOWEADLH-----LDDE 505
 Db 267 ASLTYPSCCAF-----AMNRQISDHLPCNKSLRQEVDMTOARGRVSLAEDG 318
 QY 506 ESSKRPLGLLARAENHYDODLDELQLEMEDSKHPVSPGPPKPCPELFESGIR 565
 Db 319 ESS-----LAKEDTYMSEDDLCNEVD-----VISPEDPTNPECIDGHDLIR 366
 QY 566 LAWAIVLISVLCNGVLTIVFAGGAPLPPVKEVGAIGANTLTGISCGLASVDALT 625
 Db 367 VILMFISILATIGN-ILIVILITISQYKLLPRLMCLNLAFAADLCIGIILLINSVDHT 425
 QY 626 FQGFSEYGAEMWGLGRANGFLAVLGSESVLLTLTA-----VQCSVSYSC 673
 Db 426 KTOYHNVAIDMOTGAGCDAGFFVFASELSVYTLTITLERMTITHAMOLOCKV---- 481
 QY 674 VRAVGRSPISGVSRAVGLGCLALAGLAALPLASVGEYSGASPLCTPYAPREGOPALGET 733
 Db 482 -----QLRHAASIMLVGNITAFYVALPPIFGISSTMKYSLCPM--DIDSPLSQLYV 531
 QY 734 VALVMNNSFCFLVAVGAYIKYCDLPRGD-FAVWDCAMVRHVAMLIJFADGLCYPAVFL 792
 Db 532 VSLILVNLVAFVVICGCTHTYLTVRNPNIMSSSDTKIAKRAMMLIFTDLGMAPISEF 591
 QY 793 SFASMLGLFPVTPRAVAVSVLLVYPLPACLNPILYLFPNPFRRDLRLRBRAG 846
 Db 592 AISASLKVPLITVSKSKILLVLFYVPCANPFLYAIFTKNFRDVFILLSKFG 645
 RESULT 6
 FSHR_RAT
 ID FSHR_RAT STANDARD: PRT: 692 AA.
 AC P20395;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Follicleprolin receptor).
 GN FSHR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Sertoli cells;
 RX MEDLINE=91125358; PubMed=2126341;
 RA Sprengel R., Braun T., Nikolics K., Segaloff D.L., Seeburg P.H.;
 RT "The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA.";
 RT Mol. Endocrinol. 4:525-530(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92149579; PubMed=1738373;
 RA Heckert L.L., Daley I.J., Griswold M.D.;
 RT "Structural organization of the follicle-stimulating hormone receptor gene.";
 RT Mol. Endocrinol. 6:70-80(1992).
 CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: SERTOLI CELLS AND OVARIAN GRANULOSA CELLS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. FSH/LSH/TSH SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 6 LECICINE-RICH REPEATS (LRR).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 Db EMBL: L02842; AAA41175.1; -.
 Dr PIR: A34548; A34548.
 Dr PIR: A41729; A41729.
 Dr HSSP: P23945; 1XUN.
 Dr GCRDB: GCR_0234; -.
 Dr GCRDB: GCR_0456; -.
 Dr InterPro: IPR000276; GPCR_Rhodpsn.
 Dr InterPro: IPR001611; LRR.
 Dr InterPro: IPR000372; LRR_Nterm.
 Dr Pfam: PF000560; LRR_3.
 Dr Pfam: PF01462; LRRNT; 1.
 Dr PRINTS: PR00373; GLYCHROMER.
 Dr PRINTS: PR01143; FSHRECEPTOR.
 Dr SMART: SM00013; LRRNT; 1.
 Dr PROSITE: PS00237; G_PROTEIN_RECEP_FL_1; 1.
 Dr PROSITE: PS0262; G_PROTEIN_RECEP_FL_2; 1.
 Dr G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 692
 FT DOMAIN 18 365
 FT TRANSMEM 366 386
 FT DOMAIN 387 397
 FT TRANSMEM 398 420
 FT DOMAIN 421 442
 FT TRANSMEM 443 464
 FT DOMAIN 465 484
 FT TRANSMEM 485 507
 FT DOMAIN 508 527
 FT TRANSMEM 528 549
 FT DOMAIN 550 572
 FT TRANSMEM 573 596
 FT DOMAIN 597 607
 FT TRANSMEM 608 629
 FT DOMAIN 630 692
 FT REPEAT 44 68
 FT REPEAT 69 93
 FT REPEAT 119 143
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 218 240
 FT DISULFID 441 516
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 SQ SEQUENCE 692 AA; 77681 MW; 267EA78C7CFD8EC6 CRC64;
 Query Match 10.2%; Score 518.5; DB 1; Length 692;
 Best Local Similarity 23.8%; Pred. No. 3.8e-28;
 Matches 201; Conservative 123; Mismatches 304; Indels 217; Gaps 30;
 QY 115 SIKITMLQNNQUGIPAEIMWELP-SLOSRLDANLISLIPRSPEGLSTLHMLWDNA 173
 Db 26 SIKVFLQDSKTEITPT-----DLPRNALELRVLTKRYIPGSGAGGDLKEIKISQND 81
 QY 174 LFE-IPVARNLNPALQAMTL-ALNRISHIDYAFONTLSVYLHLHNNRQHLGTHSFE 231
 Db 82 VLEVTEADVFSNLPKLEIRIKANNLLYINPEARONPLSLKYLISNTGIRKL-----P 136
 QY 232 GHANLETDLNTKNOEPPVAIRITGRLOELGCFHNNNKAIPKAPMGNPILQTHFYDN 291
 Db 137 AVHKIOSIQ-----KVLIDIQD-----NINIHIVARNSEMGIS-----FES 172
 QY 292 PLOVGRSAFOYLPKILHTLSNGAMDIOEPDLKGTSLSEILTLRAGIRLPSGMCOOL 351
 Db 173 VILMISKNGIE--EIHNCAPNG-----TQDELNTLS----- 201
 QY 352 PRLRVLELSHNOIEELPSLHRCQKLEIGLQHNRIWEIGADTFQSLSLQALDLSMNAIR 411

```

Db 202 -----DNNMLDELPN-----DVFQASGPVILDISFRKVH 231
Qy 412 SIHPEAFSTLHSLVKLDLNDQTLPLAGLGLMLKLGKLNALSGAFSEKDSFRKRL 471
Db 232 S-----LPNHGLENLKLRLARSTYRLKLPVLDKFEVTLMEA 267
Qy 472 EYVYAYOCCPYGMCASFRRASGOWEAEDLH-----LDDEE 506
Db 268 SLTYSHCCAF-----ANLKRQISELHPICNKSILRODIDDMQIGDQRSLIDDEP 319
Qy 507 SSKRPLGLLARAENHYDDDELQLEMEDSKHPVSQSPTRPPKPCYEYLFESWGIRL 566
Db 320 S---YGGSGDMMYNEFDYLCN---EYVD-----VTCSPKPDANPCEDINGYNILRV 366
Qy 567 AVMAIVLLSVLCNGVILTVLVFAGGAPRLPPVKFVGAIGANLTGSCGLASVALTF 626
Db 367 LIMEFSLIATIGNTVLY-VLTTSQYKLTVPRLMGNLAFADLCIGYLLILASVDHTK 425
Qy 627 GOFSEYGARWETGLGCRATGFLAVLGSSEASVLLTLAAVQ-----CSVSVSCVRAYGKS 680
Db 426 SOYHNYAIDMOTGAGCDAGGFTVFASLSVTLTITLERMHTITHAMQLEC-----KV 480
Qy 681 PSLGVRAGVGLCLAGLAAALPLASVGEYCASPLCLPYAPPEGPALGFTVALVMAN 740
Db 481 QLRHAASVAVLG-WTFEFAALFPIFGISSYKVSICLPM--DIDSPLSQLYVMALVLN 537
Qy 741 SFCEFLVAGAVIKLYKCDLPRGDF-EAWMDCAVRVHVMILFADGLLYCPAFELSPASMLG 799
Db 538 VLAFFVIGCGYHILYLVANPTVSSSPTKAKRNATILFTDFLMAPIFSFPAISLSK 597
Qy 800 LPPVTPAVKSVLLVLPPLACINPLLYLFPNPFEDDLRLRPRAAGSGPLAYAAAGEL 859
Db 598 VPLTVSRKAKILLVLEYPINSCANPFLYAFITKFNRRDFEIL-----L 640
Qy 860 EKSSCDSTQALVAFEDVDLLEASBAGRPGLTYGFPSVTILSCQPPAPRELSHCVE 919
Db 641 SKFGYEMQAOIYRT-----ETSSA-----THNF-HARKSHCS-SAPRVTSYVLYV 684
Qy 920 PEGNH 924
Db 685 PL-NH 688

```

RESULT 7

FSHR_HUMAN STANDARD; PRT; 695 AA.

AC P33945;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin receptor).

GN FSHR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RX MEDLINE=91222171; PubMed=1709010;

RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;

RT "Cloning and sequencing of human FSH receptor cDNA."

RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=93246012; PubMed=1301382;

RA Keiton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,

RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzaja J.B., Luchette C.A.,

RA Chappel S.C.;

RT "The cloning of the human follicle stimulating hormone receptor and

```

RT its expression in COS-7, CHO, and Y-1 cells.";
RL Mol. Cell. Endocrinol. 89:141-151(1992).
RL [3]
RP SEQUENCE FROM N.A.
RA Tilly L.T., Alhara T., Nishimori K., Jai X.-C., Billig H.,
RA Kowalski K.I., Perlas E.A., Hsueh A.J.;
RL Submitted (xxx-1992) to the EMBL/Genbank/DBJ databases.
RL [4]
RP SEQUENCE OF 1-342 FROM N.A.
RC TISSUE=Testis; PubMed=1359889;
RX MEDLINE=93075197; PubMed=1359889;
RA Gromoll J., Gudermann T., Nieschlag E.;
RT "Molecular cloning of a truncated isoform of the human follicle
RT stimulating hormone receptor.";
RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).
RL [5]
RP SEQUENCE OF 1-51 FROM N.A.
RX MEDLINE=95011044; PubMed=7926278;
RA Gromoll J., Dankbar B., Gudermann T.;
RT "Characterization of the 5' flanking region of the human follicle-
RT stimulating hormone receptor gene.";
RL Mol. Cell. Endocrinol. 102:93-102(1994).
RL [6]
RP 3D-STRUCTURE MODELING OF 49-228.
RX MEDLINE=96363672; PubMed=8747461;
RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,
RA Handrickson W.A., el Tayar N.;
RT "Structural predictions for the ligand-binding region of glycoprotein
RT hormone receptors and the nature of hormone-receptor interactions.";
RL Structure 3:1341-1353(1995).
CC -1- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: SEROTOL CELLS AND OVARIAN GRANULOSA CELLS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 5 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M65085; AAA52477.1; -
DR EMBL; S59900; AAB26480.1; -
DR EMBL; M95489; AAA52478.1; -
DR EMBL; X68044; CAA48179.1; -
DR EMBL; S73199; AAB32071.1; -
DR PIR; JN0122; JN0122.
DR PIR; JN0122; JN0122.
DR PIR; JN0122; JN0122.
DR GCRDB; GCR 0071; -
DR GCRDB; GCR 0404; -
DR GCRDB; GCR 0588; -
DR GCRDB; GCR 0690; -
DR MIM; 136435; -
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCHROMONER.
DR PRINTS; PR01143; FSHRECEPTOR.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW

```

KW Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;
 3D-structure. 1 17
 FT STGNAL 18 695
 FT CHAIN 18 366
 FT DOMAIN 18 367
 FT TRASMEN 367 387
 FT TRASMEN 388 398
 FT TRASMEN 398 421
 FT TRASMEN 422 443
 FT TRASMEN 444 465
 FT TRASMEN 466 485
 FT TRASMEN 486 508
 FT TRASMEN 509 528
 FT TRASMEN 529 550
 FT TRASMEN 551 573
 FT TRASMEN 574 597
 FT TRASMEN 598 608
 FT TRASMEN 609 630
 FT TRASMEN 631 695
 FT TRASMEN 695 93
 FT REPEAT 119 143
 FT REPEAT 143 192
 FT REPEAT 192 216
 FT REPEAT 216 240
 FT REPEAT 240 257
 FT DISULFID 442 517
 FT CARBOHYD 191 191
 FT CARBOHYD 199 199
 FT CARBOHYD 293 293
 FT CARBOHYD 318 318
 FT CARBOHYD 224 285
 FT VARSPLIC 342 695
 FT CONFLICT 13 13
 FT CONFLICT 112 112
 FT CONFLICT 197 198
 FT CONFLICT 295 295
 FT CONFLICT 307 307
 FT CONFLICT 680 680
 SQ SEQUENCE 695 AA; 78294 MW; 723B8E71F76D2CD5 CRC64;
 Query Match 10.28; Score 518; DB 1; Length 695;
 Best Local Similarity 22.28; Pred. No. 4.1e-28;
 Matches 189; Conservative 127; Mismatches 300; Indels 236; Gaps 25;

39 CHCEDGIMLSADCEGLSAVPGDLPLATYALDLSMNNTLELPGLFPHHLELRIS 98
 23 CHCSNRPL-----COESVTRIPSDLPRLNALELFVLKLVKIQK----- 64
 99 GNLHSHIPGQAFGLYSL-KIIMLQNNOLGIPAEALWELPSLOSRLDA--NLISLVE 155
 65 -----AFSGFGDLEKTEISQNDVLEADVFNSLPKLEHRIEKEKANNLLYINPE 114
 156 RSFEGSLSLRLMLDNLATLEIPRALNNLPALQAMTALNRISHIPRYAONLISVLV 215
 115 -----AFQNLPLNLYLLSLNSGKIKLP----- 137
 216 HLHNNRIQHLGTHSEGHANLETLDLNANKLOEFPVALRTGLRLOELFHNHNNIKAIPEK 275
 138 -VHK-----IHSIQKVLDD-----ID-----NINHTIERN 163
 276 AFPMGNPLQTHFYDNPIDQVGRSAFYQLPKLHLISLNGAMDIOEPDLKGTTSLEITL 335
 164 SFVGLS-----FESVILMLINKNGIQ-----ETHNCAFNQ-----TOLDELNL 200
 336 TRAGIRLRLPSGMCQQLPRLRLVLELSHNOIEELPSLHROCKLEIGLQHNRIWEIGADPFS 395
 201 \$-----DNNNIEELPN-----DVPH 215
 396 QLSLQALDLSWNAIRSIHPEAFSTHSLVKLDLTDNDOTLLPLAGLGIMHLKLGKLNLA 455
 216 GASGVYIIDISTRTHS-----LPYGLWELMKIKLARSTVN 251
 456 LSAQSKSPFKRLLEVPYAVQCCPYGMCASFASQGWEMEDLH-----LDD 504

DB 252 LKKLPLEKLVALMEASLTVPSHCCAF-----ANWRQISELHPICKNSILIRQEVY 303
 QY 505 EESSKRPLGLLAROAEHNYDODDELQLEMEDSKPH--SVQSPTRGPCKCEYLFESEW 562
 DB 304 MTOTRGROSSLAEDNESYSRGPDMTYTFEDYDLCNEVVDYCSKPKDANPCDDINGYN 363
 QY 563 GIRLAWAIVLLSVLNGVLVLFVAGGPAPLPVVKFVGAIGANTLTGISCGLASVD 622
 DB 364 ILRLVIFISILAITGN-IIVLVITTSQYKLVPRFLMCMIAFADICIGIYLLIASVD 422
 QY 623 ALTFGQSESGARWETGLGRATGFLAVLGEASVLLTLTAAG-----CSVSVQVRA 676
 DB 423 IHTSQQHNVAIDMOTAGGCDAAEFTVFASELSVYTLTATTLERWHTITHAQDQ--- 479
 QY 677 YGRSPISGSVAVGLGCLALAGLAAALPLASVGEYSGASPLCLPYAPGEGOPALGFTVAL 736
 DB 480 --KQLRHAASVVMGMV--PAPALPPIFGISSMYVSIQLP--DIDPSLQLYVMSL 534
 QY 737 VMNNSFCELVVAGVYIKLYCDLPRGDF-EAVWDCAMVRHVAWLIFADGLYCPVAFLSFA 795
 DB 535 LVNLVAFVVICGYIHLYTVRNPNTYSSSDPRIRAKRMAMLIETDFLCMAPISFFAIS 594
 QY 796 SMLGLFVTPRAVKSVLVYVLPACLNPLLYLTFNPHFRDRLRLRPAGDSGLAYAA 855
 DB 595 ASLKVPLTVSKAKILVLFHPINSCANPFLYALFTKNFRDFFILSKG-----CYEM 649
 QY 856 AGELEKSSCDST 867
 DB 650 QAOIYRTETST 661

RESULT 8
 FSHR_MACFA ID FSHR_MACFA STANDARD; PRT; 695 AA.
 AC P32212;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin receptor).
 GN FSHR.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecinae; Macaca.
 CX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=94071854; PubMed=7504463;
 RA Gromoll J., Dankbar B., Sharma R.S., Nieschlag E.;
 RT "Molecular cloning of the testicular follicle stimulating hormone receptor of the non human primate Macaca fascicularis and identification of multiple transcripts in the testis";
 RT Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
 CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/FSH SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: X74454; CAA52463.1; -.
 DR PIR: S36452; S36452.

```

DR PIR; JN0898; JN0898.
DR HSSP; P23945; 1XUN.
DR GCRDb; GCR_0653; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 4.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PRO0373; GLYCHOMONER.
DR PRINTS; PRO1143; FSHRECEPTOR.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17
FT CHAIN 18 695
FT DOMAIN 18 366
FT TRANSSEM 367 387
FT TRANSSEM 388 398
FT TRANSSEM 399 421
FT DOMAIN 422 443
FT TRANSSEM 444 465
FT TRANSSEM 466 485
FT TRANSSEM 486 508
FT DOMAIN 509 528
FT TRANSSEM 529 550
FT DOMAIN 551 573
FT TRANSSEM 574 597
FT DOMAIN 598 608
FT TRANSSEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT DISULFID 218 240
FT CARBOHYD 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
SQ SEQUENCE 695 AA; 78343 MM; OD60A233729B5250 CRC64;

Query Match 10.2%; Score 515; DB 1; Length 695;
Best Local Similarity 22.7%; Pred. No. 6.6e-28;
Matches 195; Conservative 117; Mismatches 298; Indels 248; Gaps 26;

```

```

QY 396 QLSIQALDLSWMAIRSHHPAFTLSLVLKIDLTDMQTLPLAGLGMLKIKGNIA 455
DB 216 GASPEVLIDISRTIRHS-----LPSYGLNLEKIKRARSTYN 251
QY 456 LSCAFKDSFPEKRLIEPVAYVOCPCYMGCAFSPKASGOMEAEDH-----LDD 504
DB 252 LKRLPSLEKVALMEASILTPSHCCAF-----ANMRROISELHPICNNSILKEQVY 303
QY 505 EESKRPGLGLARQAEHNYDDDELQLEMEDSKRP--SVQCSPTPGFPKCEYLFRSW 562
DB 304 MTQRGORSSLAEDNESSYRGFMTRYAEFDYDLCNEVVDYVTCSPKPAFNCEDILGVN 363
QY 563 GIRAAVAIVLISVGLNGVILVTFAGAPRALPVKFPVGAAGANTLTGSCGLIASVD 622
DB 364 LKRLVIFISLATTGNIVLVTL-TTSQYKLTVPRLMCNLAFAADLCIGYLLIASVD 422
QY 623 ALAFGPSEYGARWETGLGRATGFLAVLGSSEAVLITLTAVQ-----CSVS 670
DB 423 IHTKSQYHNVAIDMQTGAGDAAGFFVFPASELSVYTLTALTLEMTHTITAMODCKVH 482
QY 671 VSCVRATGKPSLGSVAVYGLCLAGLAADPLASVGEYGSAPLCLVPAPPEGOPAL 730
DB 483 VR-----HAASVMWGMV-FAPAAALFPFGISSYMKVSIQCPM--DIDPSLQ 528
QY 731 GFTVALVMNSFCFLVYVAGAYIKLYCDLPBGDF-EAWMDCAVHVHAMLIAADGLYCPV 789
DB 529 LYVMSLVLVNLAFAVVICGCTHYTLVVRNPNIYSSSDRIARMAWLITFDLCMAPI 588
QY 790 AFLSPASMLGLFPVTPAVKSVLLVLPACLNPLLYLTFNPFRRDLRLRAPAGDSG 849
DB 589 SFPAISASLVPVLTISKAKILLVFPYNSCAMPFLYAIPTKKNRDFLILKFG--- 645
QY 850 PLVYAAAGELEKSSCOST 867
DB 646 --CYEMQAQIYKRTSTST 661

RESULT 9
FSHR_EOUCAS STANDARD; PRT; 687 AA.
AC 095179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Folicle stimulating hormone receptor precursor (FSH-R) (Folliotropin
DE receptor).
GN FSHR.
OS Equus asinus (Donkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9793;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis.
RX MEDLINE=97338913; PubMed=9195473;
RA Richard F., Martinat N., Remy J.-J., Salesse R., Combarrous Y.;
RT "Cloning, sequencing and in vitro functional expression of
RT recombinant donkey follicle-stimulating hormone receptor: a new
RT insight into the binding specificity of gonadotropin receptors.";
RL J. Mol. Endocrinol. 18:193-202(1997).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/FSH SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```

CC -1- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17146; CA35026.1; -
DR EMBL: X17147; CA35027.1; -
DR EMBL: M29957; AAA30901.1; -
DR EMBL: M90047; AAA30902.1; -
DR PIR: S06933; S06933.
DR PIR: A40077; A40077.
DR HSSP: P16473; 1XDM.
DR GCRDB: GCR_0020; -
DR GCRDB: GCR_0021; -
DR GCRDB: GCR_0022; -
DR GCRDB: GCR_0224; -
DR InterPro: IPRO00276; GPCR_Rhodopsn.
DR InterPro: IPRO01611; LRR.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR; 3.
DR PRINTS: PR00373; GLYCHROMER.
DR PRINTS: PR01145; TSHRECEPTOR.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECPT_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing.
FT SIGNAL 1 20
FT CHAIN 21 764
FT DOMAIN 21 413
FT TRANSMEM 414 441
FT DOMAIN 442 450
FT TRANSMEM 451 473
FT DOMAIN 474 494
FT TRANSMEM 495 517
FT DOMAIN 518 537
FT TRANSMEM 538 560
FT DOMAIN 561 580
FT TRANSMEM 581 602
FT DOMAIN 603 625
FT TRANSMEM 626 649
FT DOMAIN 650 660
FT TRANSMEM 661 682
FT DOMAIN 683 764
FT REPEAT 51 74
FT REPEAT 150 174
FT REPEAT 176 199
FT REPEAT 201 223
FT REPEAT 223 259
FT DISULFID 494 569
FT CARBOHYD 77 77
FT CARBOHYD 99 99
FT CARBOHYD 177 177
FT CARBOHYD 198 198
FT CARBOHYD 302 302
FT VARSPIC 81 105
SQ SEQUENCE 764 AA; 86483 MW; 49F03B3DBCB65512 CRC64;

Query Match 9.7%; Score 490.5; DB 1; Length 764;
Best Local Similarity 22.7%; Pred. No. 3; 7e-26;
Matches 211; Conservative 131; Mismatches 353; Indels 233; Gaps 30;

```

```

QY 123 NNOIGTIPAEALWEIPSLQSLDANLISLVPERSEGISLRHML-DONALTEIPVRA 181
DB 85 -----IDATLQRL-ESHSPYNLSKMTHEIRTRSLTSDPDA 121
QY 182 LNNLPALQAMTLALNISHIPDY-----AFQNL-TSLV 214
DB 122 LKELPLKRLKGLFNPNGLGFPVTVYSTDVFFLEITDNPMASTIPANAFGLCNETYT 181
QY 215 LHLHNNRIQHLGTHSEFEGHNLLETDLNKNLQEPVVALRTGLQELGFHNNIKAIPE 274
DB 182 LKLYNNGFTSIQGHAFNGT-KIDAVLYLNKKY-----LSAIDX 218
QY 275 KAFMGNPILQTHFYDNP-----IQFVGNSAF--QYLPKHLTSLGAMDIOEPDLKGT 328
DB 219 DAFGG-----VSGPFLDVSTYSVALPSKLEHLEKELIARNTWLKRLP-----L 265
QY 329 SLEILTLTRAGIRLPSGMCQOLPRLRYVLELSHNGIEELPSLRQCKLEIGLQNNRIWE 388
DB 266 SLSPFLHTRADLS-YPSHCC-----AFKNQKTRGT----- 295
QY 389 IGADTFQSLSLQALDLSMNAIRSIHP-EAVSTLSLVKLDLTDNQTLPLAGGLMH 447
DB 296 -----LESIMCNESIRSLRQKRSVTLNGPPEQEEYVL-----GDSH 334
QY 448 LKLGKLNALSGAFSKDSPKRLILEVPYAYOCCPYGMCASFEGKASGWEADLHIDDES 507
DB 335 AGYKDN---SQFQDTDSHSHYV-----FPE-----EQDEILIGFGQE 369
QY 508 SKRPLGLARQAEHNYDODLDELQLEMEDSKPHSVQSPPTPGPPCEYLFESWGIRLA 567
DB 370 LKNPQETLQAFDSHYDYV-----CGNEMWCTPSDENPCEDLMGYKFLRIY 420
QY 568 VMAIVLTVSLONGVLYLVFAGGAPRLPYKVFYVGAIGANTLIGISGLIASVDALTFG 627
DB 421 VWEVSLTLGNVLYLVLSHKKLYVP-RFLKCNLAFAFCMGWMTLLTSLASVDLYTHS 479
QY 628 QFSEYGAEMETGLCARATGFLAVLSEASVLLTLAAVQCSVSVYARVAKSPSLGVR 687
DB 480 EYVHAIIDMOWGPCNAGFTTFVASELSVTLVITLIERKATITFAMRLDKRLRIAY 539
QY 688 AGVIGCLALAGLAALPLASVGEYASPLCLPYAPREGOPALGFTVALVMNSFCELV 747
DB 540 AINMGWCCFLMLPLVIGISYAKVSIQPM--DFETPALAVIILVLLNIVAFIIV 597
QY 748 AGATYIKLYCDLPRDDE-AWMDCAVHRVAMLIRADGLLYCPVAFSLASMLGEPYIPE 806
DB 598 CSCYVKIITYIRNQNYPNGDKDTIARMAVLIITDFCMAPISFYALSALMNRPLITVT 657
QY 807 AVKSVLLVLPPLACLPNPLYLIFNPHFRDRLRLRPAGDSGLAYAAAG----- 857
DB 658 NSKILLVLEVPDNCANPDLAIFTKAFQORVFIILSKFGICRQQAAYRQGRVPSKNSA 717
QY 858 --ELEKSSCSTQALVAFSDVDLILEAS 883
DB 718 GIQIQKVTBDMROSLPMMODEYELLENS 745

RESULT 11
TSHR_MOUSE
ID TSHR_MOUSE STANDARD; PRT; 764 AA.
AC P47750;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone
DE receptor).
GN TSHR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.

```


RC STRAIN-BALB/C; TISSUE=Thyroid;
RX MEDLINE=9422423; PubMed=8170469;
RA Stein S.A., Oates E.L., Hall C.R., Grumbles R.M., Fernandez L.M.,
RA Taylor N.A., Puett D., Jin S.;
RT "Identification of a point mutation in the thyrotropin receptor of
the hyl/hyl hypothyroid mouse";
RL Mol. Endocrinol. 8:129-138(1994).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DISEASE: DEFECTS IN TSHR ARE THE CAUSE OF HYPOTHYROIDISM.
CC AN AUTOSOMAL RECESSIVE, FETAL-ONSET, SEVERE HYPOTHYROIDISM RELATED
TO TSH HYPORESponsiveness AND ASSOCIATED WITH ELEVATED TSH.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 8 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: 002601; AAA53209.1; -
CC EMBL: 002602; AAB0455.1; -
CC HSSP: P16473; IXUM.
CC GCRdb: GCR_1718; -
CC GCRdb: GCR_1719; -
CC MGD: MGI:98849; Tsh.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC InterPro: IPR001611; LRR.
CC Pfam: PF00001; 7tm_1; 1.
CC Pfam: PF00560; LRR; 3.
CC PRINTS: PR00373; GLYCHOMONER.
CC PRINTS: PR01145; TSHRECEPTOR.
CC PROSITE: PS00237; G-PROTEIN_RECEP_FL1; 1.
CC PROSITE: PS00262; G-PROTEIN_RECEP_FL2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Phosphorylation; Repeat; Leucine-rich repeat; Disease mutation.
CC SIGNL. 1 21
CC CHAIN 22 764
CC DOMAIN 22 413
CC TRANSMEM 414 441
CC DOMAIN 442 450
CC TRANSMEM 451 473
CC DOMAIN 474 494
CC TRANSMEM 495 517
CC DOMAIN 518 537
CC TRANSMEM 538 560
CC DOMAIN 561 580
CC TRANSMEM 581 602
CC DOMAIN 603 625
CC TRANSMEM 626 649
CC DOMAIN 650 660
CC TRANSMEM 661 682
CC DOMAIN 683 764
CC REPEAT 51 74
CC REPEAT 100 124
CC REPEAT 125 150
CC REPEAT 151 174
CC REPEAT 176 199
CC REPEAT 200 223
CC REPEAT 225 248
CC REPEAT 264 288
CC DISULFID 494 569
CC CARBOHYD 77 99
CC CARBOHYD 99 99
CC CARBOHYD 177 177
CC CARBOHYD 198 198

FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 556 556 P -> L (IN HYPOTHYROIDISM).
SQ SEQUENCE 764 AA; 86583 MW; 6ADD2CC72P018317 CRC64;
Query Match 9.6%; Score 486; DB 1; Length 764;
Best Local Similarity 21.7%; Pred. No. 7.6e-26;
Matches 211; Conservative 132; Mismatches 346; Indels 284; Gaps 30;
QY 4 PGLRLMLCALACARRAGAPQPGPAPACPHC-QEDGIMLSADCSGLISAVPG 62
DB 3 PGLLLLVLLALSRSLR-----GKECAPPECHQEDDFRT--CKEL----- 44
QY 63 DLDPLTAYLDLSMNLTEPLGFLHRLRLEIRLSGNLSHISPGAFSGLYSLKIMQ 122
DB 45 -----HRLPSLPS-----TQTLKLETHLKTPLSLAFSLPNIISRTYS 84
QY 123 NNOIGCPALMELPSLOLRDANMLISLPRSPREGSLSLH-----WLDNA 173
DB 85 -----IDATLQRLP--HSFYNSKMHIELRMTSLFTYIDPA 121
QY 174 LTEIP-----VRALNLPALQAMTLAL-----NRISHIDYAFQNL-TSLV 214
DB 122 LTELPLKFLGIFNTGIRIPDLTKIYSTDIFLEITDNPYTSVENAFQGLCNETLL 181
QY 215 LHLHNNRIQHLGTHSEGLNLETLNLYKQLQEPFAITLRLQELGHHNNIKAIPE 274
DB 182 LKLYNNGFTSYOGHAFNGT-KLDVAVYLNNKRY-----LRAIDN 218
QY 275 KAFNGNLLQTHFYDNP-IOEFGSAFOYLPR-----LHLSINGAMDIOEPDLKGT 328
DB 219 DAFGG-----YSGFTLLDVSTSVTALPSGHEHLKELAKDWTLKLP-----L 265
QY 329 SLEILITLRAGIRLLPSGMCQQLRLRYLELSHNOIELDPSLHRCQKLEETIGLOHRIWE 388
DB 266 SLSPFLHLTRADLS-YPSHCACFKNQKIRIGLESIMKNESIRNLORKSVNLLGPIYO 324
QY 369 -----IGAPTFQGLSSICALDLSWMAIRSHPEAFSTLSLYKLDLTDLNQLTLP 438
DB 325 EYEDPDGDSVYG---YKNSKFKQ-----ESPNSHIYY----- 354
QY 439 IAGLGCLMHLKLGMLSLQAFSKDSEPKIRILEVPYAYOCPPGMCASFPGKSGOMEAE 498
DB 355 -----FEE-----EOE 360
QY 499 DLHIDDESSKRPGLLARQAEHNYDDDLQLEMEDSKRHPVOCSPPTGPKPCEYL 558
DB 361 DEVVFGQELKNPQRETLQAFESHYDTY-----CGNEDMVCPTPSDEFNPEEDI 411
QY 559 PESMGIRLAVNAIVLVICNGVLLVTFAGGAPALPVPKVVGAIGANTLTISGGL 618
DB 412 MGYRFLRLVWFVSLALLGNIFFVLLILLSHYKLTVE-RLMGNLAFADFCMGVYLLLI 470
QY 619 ASVDALITGQFSEYGAKEWTGICRATGFLAVLGSSEASVLLITLAOVCSVSCVARYG 678
DB 471 ASVDLYTHSEYNNHAIIDWOTGPGCTAGFFVFVASELSVYLLTVTLTRWATTFPAMRLD 530
QY 679 KSPSLGSVAVAGVGLALAGLAALPLASVGEYSAPCLPAPPEGPPALAGFTALVM 738
DB 531 KRIRLRHAYTTIAGWVSCFTLALLPMWGISYAVSICLPR--DTDFPLALAVYVLL 588
QY 739 MNSFCFLVAGAVYIKLYCDL-----PRGDFEAVWDCAVNRHVAVLIPADGLITVPVAF 792
DB 589 LNVAVFVVCSCYVKIITYVRNPQYNPRDK-----DTKIAKRAVALIFTDFECMAPISEY 643
QY 793 SFASMLGLFPVPEAVKSVLLVYLPACNLPLVILFPHPRDRLRRIRPAPGDSGLA 852
DB 644 ALSALNNKPLITYVNSKILLVLEPLNSCANPPLTALFTKAFQKQVYFILLSFGICKRQA 703
QY 853 YAAAG-----ELKSSCDSTQALVAFSDVLLILEASEAGRPGLFTTYGPPSVTL 901
DB 704 QAYQGRVCPNNSTGIQIKITPDTRQSLPNNQDYTELLGNSQL----- 747
QY 902 ISCOGPGAPRLGEG 914

```

Db 748 -----APKLOG 753

RESULT 12
TSR-BOVIN STANDARD: PRT: 763 AA.
AC Q2/987;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thymotrophin receptor precursor (TSR-R) (Thyroid stimulating hormone
DE receptor).
CN TSR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RA MEDLINE=97280128; PubMed=9134497;
RA Silverides D.W., Houde A., Ethier J.F., Lussier J.G.;
RT "Bovine thymotrophin receptor cDNA is characterized by full-length and
RT truncated transcripts."
RL J. Mol. Endocrinol. 18:101-112(1997).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
CC CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
CC CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSR SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U15570; AAC38639.1; -
CC HSSP: P16473; 1XDM.
CC
CC GCRDB: GCR_1225; -
CC InterPro: IPR000276; GPCR_Rhodopsin.
CC InterPro: IPR001611; LRR.
CC Pfam: PF00001; 7tm.1; 1.
CC Pfam: PF00560; LRR; 3.
CC PRINTS: PR00373; GLYCHROMONER.
CC PRINTS: PR01145; TSRRECEPTOR.
CC PROSITE: PS00237; G_PROTEIN_REC_P1.1; 1.
CC PROSITE: PS50262; G_PROTEIN_REC_P1.2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 21
FT CHAIN 1 763
FT DOMAIN 22 412 THYROTROPIN RECEPTOR.
FT TRANSMEM 413 440 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 441 449 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 450 472 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 473 493 2 (POTENTIAL).
FT TRANSMEM 494 516 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 517 536 3 (POTENTIAL).
FT TRANSMEM 537 559 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 560 579 4 (POTENTIAL).
FT TRANSMEM 580 601 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 602 624 5 (POTENTIAL).
FT TRANSMEM 625 648 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 649 659 6 (POTENTIAL).
FT TRANSMEM 660 681 EXTRACELLULAR (POTENTIAL).
FT 7 (POTENTIAL).

```

```

FT DOMAIN 662 763 CYTOPLASMIC (POTENTIAL).
FT REPEAT 51 74 LRR 1.
FT REPEAT 125 150 LRR 2.
FT REPEAT 151 174 LRR 3.
FT REPEAT 176 199 LRR 4.
FT REPEAT 201 223 LRR 5.
FT REPEAT 225 248 LRR 6.
FT DISULFID 493 568 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 763 AA; 86431 MM; 3589647 EDTA8C CRC64;

Query Match 9.5%; Score 479; DB 1; Length 763;
Best Local Similarity 23.1%; Pred. No. 2,3e-25;
Matches 210; Conservative 139; Mismatches 321; Indels 238; Gaps 36;

QY 31 GPTACPA-PCHC-QEDGIMLSADCELSGSAVGDPLTAYLDLSMNNLTIELQGLPHH 88
DB 20 GGERCPSPCECROEDDEPRVT--CKD--IQSIP-SLPEST-----QT 56
QY 89 LRFELRLSGNHLSHIPGAFSGLYSLKILMLQNNOLGIGTAEALWELPSLQSLRDN 148
DB 57 LKFI-----THLKTISRAPSNLPNISRYLS-----IDAT 88
QY 149 LISLVERSEFGLSLRLHL--DONALTEIPYALNNLPAQAMTLALNRISHIPDAFQ 207
DB 89 LQOL-ESHSEFYNLSKVTHIEIRNTRSLTYIDSGALKEPLKFLGIFNTGLRVFPD----- 143
QY 208 NLTSL-----VVLHLNN-RIGHGTSFEGHNLFTLDLNYKNLQEPVAIRTLGRLO 260
DB 144 -LTKIYSDVFFLEITDNPWTSTIPANAFQGLCN-ETLTLC----- 183
QY 261 ELGHNNNKAIPEKAFGNPDLQTIHFYDNP-IOFVRSRFAOYLPKHLTSLGANDIQ 319
DB 184 ---LYNNGFTSIQGHAFNGCR-LDAVINNKMYKLYVIGQDAF----- 221
QY 320 EFPDLKGTSLTEILTLRAGIRLLPSGMCQQLPRLRYLEISHNOIEELPSLRQKLEEI 379
DB 222 -----AGVYSGPT-----LSDISYTVTALPSK----- 244
QY 380 GLQHNRIWEIGADTFQSLSLQALDLSMNAIRSIHPEAFSLHLYKDLTF-----D 431
DB 245 GLEH-----LKEILARN-TW-TLRKL-PLSLSLH-LTRADLSYSHCCAFK 287
QY 432 NQLTPLPLAGLGLMLKLGKGNLALSOAFKSPKLR-----ILEVPAVQCCPYGMC 485
DB 288 NO-----KKIRGIQ-SIMCNESSIRGLRKRKSALNGPFYQVEDIGDG 332
QY 486 ASPEKASGQW-----EADLHLDEESSKRPGLGLARAEHNYDODLELO 531
DB 333 SAGYKENSKEQDQNSHYVFEEOEDETLIGQOQLKNQOETLQAFDDHYDTVGGS 392
QY 532 LEMEDSKPHBSVOCSPTPGPKPCOEVLFEWSGIRLAWAIVLTVLCNGVLLTVFAGP 591
DB 393 EDM-----VCTPKSDFNFCEDIMGYKFLRIYVWFPSLALAGNVVLYILTSY 443
QY 592 APPLPYKVVYGAAGANTLTGSCGLASVDALTFGOFSEYGARWETGLGCRATFLAVL 651
DB 444 KLTVP-RFLMCNLAFDFCMGLVLLTASVDLTQSEYYNHAIDMOTGPCNAGFEYV 502
QY 652 GSEASVLLTLAIVQCSVSVCYRAYGKSPISGVSAGVAGLALAGLAAVLAAGEV 711
DB 503 ASBLAYTTLVTLLEHMAHTTPAMRDKIRIMHAYVIMGVCVCFLLALPLVGISST 562
QY 712 GASPLCLPYAPPEGOPALGFTVALVMNMSFCELVVAGAYIKLYCDLP-----GDFEAY 766
DB 563 AKVSIOLPM--DTEPLALAYIIVLLNLIATFIIVCACYKVIYIVRPNHVPD----- 616
QY 767 WDCAMVRYHVAWLIIFDGLLYCPAFLSFASMGLPFVTEAKSVLLVVLPLPACINPLL 826

```


DR	EMBL	S40787	AAB22684.2	JOINED
DR	EMBL	S40903	AAB22684.2	JOINED
DR	EMBL	S40904	AAB22684.2	JOINED
DR	EMBL	S40905	AAB22684.2	JOINED
DR	EMBL	S40909	AAB22684.2	JOINED
DR	EMBL	S40918	AAB22684.2	JOINED
DR	EMBL	S40920	AAB22684.2	JOINED
DR	EMBL	S40795	AAB22684.2	JOINED
DR	EMBL	S40798	AAB22684.2	JOINED
DR	EMBL	M68928	AAA41529.1	JOINED
DR	EMBL	M68917	AAA41529.1	JOINED
DR	EMBL	M68918	AAA41529.1	JOINED
DR	EMBL	M68919	AAA41529.1	JOINED
DR	EMBL	M68920	AAA41529.1	JOINED
DR	EMBL	M68921	AAA41529.1	JOINED
DR	EMBL	M68922	AAA41529.1	JOINED
DR	EMBL	M68923	AAA41529.1	JOINED
DR	EMBL	M68925	AAA41529.1	JOINED
DR	EMBL	M68926	AAA41529.1	JOINED
DR	EMBL	M68927	AAA41529.1	JOINED
DR	PIR	A32460	A32460	
DR	PIR	A41343	A41343	
DR	HSSE	P22888	1LUT.	
DR	GCRDB	GCR_0138	-	
DR	GCRDB	GCR_0139	-	
DR	GCRDB	GCR_0262	-	
DR	GCRDB	GCR_0612	-	
DR	GCRDB	GCR_0613	-	
DR	GCRDB	GCR_0614	-	
DR	InterPro	IPR000276	GPCR_Rhodopsin	
DR	InterPro	IPR001611	LRR	
DR	InterPro	IPR000372	LRR_Nterm	
DR	Pfam	PF000001	7tm_1	1
DR	Pfam	PF00560	LRR_2	
DR	PRINTS	PR00373	GLYCOPORIN	
DR	PRINTS	PR01144	LSHRECEPTOR	
DR	SMART	SM00113	LRRNT	1
DR	PROSITE	PS00237	G_PROTEIN_REC_P1.1	1
DR	PROSITE	PS50262	G_PROTEIN_REC_P1.2	1
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;			
KW	Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing			
FT	SIGNAL	1	26	
FT	CHAIN	27	700	
FT	DOMAIN	27	362	LUTROPIN-CHORIOGNADOTROPIC HORMONE RECEPTOR.
FT	TRANSSEM	363	390	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	391	399	1 (POTENTIAL).
FT	TRANSSEM	400	422	CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	423	443	2 (POTENTIAL).
FT	TRANSSEM	444	466	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	467	486	3 (POTENTIAL).
FT	DOMAIN	487	509	CYTOPLASMIC (POTENTIAL).
FT	TRANSSEM	510	529	4 (POTENTIAL).
FT	TRANSSEM	530	554	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	552	574	5 (POTENTIAL).
FT	TRANSSEM	575	598	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	599	609	6 (POTENTIAL).
FT	TRANSSEM	610	631	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	632	700	7 (POTENTIAL).
FT	REPEAT	52	75	CYTOPLASMIC (POTENTIAL).
FT	REPEAT	126	150	LRR 1.
FT	REPEAT	152	175	LRR 2.
FT	REPEAT	176	200	LRR 3.
FT	REPEAT	202	224	LRR 4.
FT	REPEAT	225	248	LRR 5.
FT	REPEAT	250	271	LRR 6.
FT	DISULFID	443	518	BY SIMILARITY.

Query Match	9.4%	Score 474;	DB 1;	Length 700;
Best Local Similarity	21.5%	Pred. No. 4.5e-25;		
Matches 203;	Conservative 130;	Mismatches 322;	Indels 288;	Gaps 30;

Db	6	PALROLIVLAVLLKPSQSLRELSG-----SRCPPCCOAPDGAH-----	46
OY	59	AVPDDDDPLAVAYDLSMNNNTLEQPGLFHRLRLLELRSLGSHLSHQAFSLYS-LK	11
Db	47	RCPC---PRAGLARSLTLYP-----VVVIPSQAFRGINEYVK	81
OY	118	ILMLONLGGIPAEALAMELPSLOSRLD--ANLISLVERSEFEGSSLRHLMDNMLT	179
Db	82	IEISQDSLSRIEAMNFNDNLINSELLIONTNLILIEP-GAFPLMPRLKLTISONTGIR	144
OY	176	EIP-VRLANLPLALQANTALN-RISHIDPAFQNLTS-LVYLHLNNRRIQHGHSFEG	233
Db	141	TLDPVTKISSSENFLELICDNLHITTINGAFNFOGNNESVTLKLYGNGFEFVOSHANG	200
OY	233	LHNLETLDLWYKNGLOEFPAVIRLGRLOGLGRHNNNIKAIRPKAMGNPLIQTHFDNP	293
Db	201	T-TLISLTKEN-----	211
OY	293	IQFVGSAPQYLPKHTLTLINGAMDIOERPDLKGTSLLEILTRFAGIRLPSGMCOOLP	352
Db	212	-----IYLEKMSHGAFOGATG-----PS-----	222
OY	353	RLRVLELSHNOGLEELPSLRQCKLEIGLOHNRKRIETGADTFSSQSLQALDSWMAIRS	412
Db	230	---TLDISSTKLOALPS-----HG-----LESIOFL-----	252
OY	413	IHPEAFSTLHSVLKDLDTNQLTTLPLAGLIGIMHLKLGKMLALSQAQSKSPFKRLILE	472
Db	253	-----IALSSYSLSKLTP-----SKKFTLSLYAT	276
OY	473	VPIAYOCPCYGCASFRRKASQGWEMEDLHDDSSKRPDLGLARQAEHN--YDQDDEL	530
Db	277	LYTSPSHCCAFRLMPK-----KQONSFSIFENFSSKQCESTYRKADNETIYSAIPEEN	328
OY	531	OLEMED-----SKPHPSVOCSPTPGPFPKPCPELFESMGIRLAVAYILSVLCNGVLLT	585
Db	329	ELSGMDYDGCSP-KTIQCAPEPAPNFCEDINGYAFRLVILINILAIFGN-LTVLF	386
OY	586	VRAGGPAPLPVKFEVVGAIAGANTLTGISCGLASVDALTFQGSSEKGAWEITGLGRAT	645
Db	387	VLLTSRYKTLTVPRFLMCNLSFADFQMGLYLLILASVDSQFGQYVYNAIDMOWGSGCGAA	446
OY	646	GFLAVLGSASVLLTLTAAVOCSSVSCVRYRAGSPSLGSRACVLCGLAAGLAALPL	705
Db	447	GFETYPASLSYTLTLVTLLEKMHITTYAVOLDKRLRLRAPIRLMGWLFSTLITMPL	506
OY	706	ASVGEYASPLCLPAPPEGPPALIGFVALVMMNSFCFLVAGAYIKLCIDLPDGFEX	765
Db	507	VGISVNMWYSICLPM--DYESTLSQVYITSLILINVAFAVYICACIRITFAVQNPDLRA	564
OY	766	V-WDCAMVRHVAWLIFADGLLYCVAALPSFASMLGFLPVPTEAVKSVLLVLPPLACLP	824
Db	565	PKDKRIKKAKMALIFLTFDTCMAISFEFALISAFAKVPILITVNSKILLYLFYVNSCAMP	624
OY	825	ILYLLFNHFHFDRLRLPRRAGDSGPLAVAAAGLEKSSCDSDQALVAFSDVLLIEASE	884
Db	625	FLYALFTAFQPDPLILSRG-----CCKRRA-----E	653
OY	885	AGRPGLFTY-----GFP-----SVTLISCOOPGAPR	911
Db	654	LYRRKEFSAYTSNCKNGFPGASKPSQATLKSLTSHVCOQPIPR	696
RESULT 14			
TSHR_SHEEP			
ID	TSHR_SHEEP	STANDARD;	PRT; 764 AA.
AC	P56495;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone receptor).		

RESULT 14		
TSHR_SHEEP	STANDARD;	PRT; 764 AA.
ID TSHR_SHEEP		
AC P56495;		
DT 15-JUL-1998	(Rel. 36, Created)	
DT 15-JUL-1998	(Rel. 36, Last sequence update)	
DT 16-OCT-2001	(Rel. 40, Last annotation update)	
DE	Thyrotropin receptor precursor (TSH-R) (Thyroid stimulating hormone receptor).	

GN TSHR.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=pituitary;
RX MEDLINE=97445147; PubMed=9299474;
RA Bockmann J., Winter C., Witkowski W., Kreutz M.R., Boeckers T.M.;
RT Cloning and expression of a brain-derived TSH receptor.;
RL Blochm. Biophys. Res. Commun. 238:173-178(1997).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN. PLAYS A CENTRAL ROLE IN
CC CONTROLLING THYROID CELL METABOLISM. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE
CC CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Y13434; CAA73846.1; -.
DR HSSP; P16473; 1XUM.
DR GCRdb; GCR_2587; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR Pfam; PF00001; 7tm1; 1.
DR Pfam; PF00560; LRR; 3.
DR PRINTS; PR00373; GLYCHROMONER.
DR PRINTS; PR01145; TSHRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECPT_FL_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
FT Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 21
FT CHAIN 1 21
FT DOMAIN 22 764
FT TRANSSEM 22 413
FT TRANSSEM 414 441
FT TRANSSEM 442 450
FT TRANSSEM 451 473
FT TRANSSEM 474 494
FT TRANSSEM 495 517
FT TRANSSEM 518 537
FT TRANSSEM 538 560
FT TRANSSEM 561 580
FT TRANSSEM 581 602
FT TRANSSEM 603 625
FT TRANSSEM 626 649
FT TRANSSEM 650 660
FT TRANSSEM 661 682
FT TRANSSEM 683 764
FT DOMAIN 683 764
FT DISULFD 494 569
FT REPEAT 51 74
FT REPEAT 125 150
FT REPEAT 152 174
FT REPEAT 176 199
FT REPEAT 201 223
FT REPEAT 223 248
FT REPEAT 248 277
FT CARBOHYD 77 99
FT CARBOHYD 99 177
FT CARBOHYD 177 198
FT CARBOHYD 198 302
FT CARBOHYD 302 302
SEQUENCE 764 AA; 86674 MW; 3242349705FEB896 CRC64;

Query Match 9.3%; Score 473.5; DB 1; Length 764;
Best Local Similarity 23.4%; Pred. No. 5,5e-25;
Matches 206; Conservative 120; Mismatches 322; Indels 231; Gaps 35;
35 CPA-PCHC-QEDGIMLSADSCSELGSAVPGDIDPLATVADLSMNNLTLELPGLFHHRL 92
||| ||| ||| : : : : : ||| : : : : :
24 CSPPECECRQEDDFRVT--CKD--IQRIP-SLPSPT-----OTLKEI 60
93 EELRLSGNLSHIIPOAGFSLKTIIMQNNQGLGIPAEALMELPSLQSLRLDNLISL 152
||| ||| ||| : : : : : ||| : : : : :
61 E-----THLKTIPRAFSNLPNISRITLS-----IDNITLOOL 92
153 VPSEFEGSLSLRHML--DDNALTEIPYALNNLPALQAMTL-----ALNRI---- 198
||| ||| ||| : : : : : ||| : : : : :
93 -ESHFYNLSKYTHIEIRNTRSLTYIDGSALEKLPDLKTLGIFNGLRVPLFYISYD 151
199 -----SHIPDYAFQNLTS-LVYLHLNHRIOHGLTSHFEGHNLFTLDLNTN 244
152 VFFILEITDNPYMTSVPAVAFQSLNETLTLKLYNNGFTSIGHFNGT-KIDAVYLNKN 210
245 KIQEPFVARTIGRQELGFHNNNIKAIPKAPMGNPILQTHFYDNP-----IQFVGRSA 300
211 KY-----LTVIDQDAFAG-----VYSGPLLDISTSVTA 240
301 F--QYLPKHLTSLNGAMDIOEFPDLKTTSELTITFRAGIRLPSGMCQQLPRLRYLE 358
241 LPSKGLHKLKLIANTWTYKLP-----LSLPLHLTRADLS-YPSHCC----- 284
359 LSHNOIELPSLHRCQK-----LEBIGLOHNRIMEIGADTFSQSLQALDLSMNAIRSI 413
285 -----AFKNQKNRIGILQSLMCNESSIWGL-----RQRKSASALN----- 319
414 HPEAFSTLSLTKLDTLQNLTLPLAGLGLMLKLGKLNLSQAFSDSPKRLILEV 473
320 -----GPTIOEYEDGD-----GSAG-----YKNSKFQDTHNSHR----- 352
474 PYAYOCCPYMCASFEGKASGOWEADLHLDDESSRRPLGLLARAENHYDDELQLE 533
353 -YV-----DQEDELITGFGELKNPQEFRLQAFDNHYDTVCGSSE 395
534 MEDSKRHSVQCSPTPPGPKCEYLFESWGIRLAWAYLVLCNGVLLVFPAGRP 593
396 M-----YCPKSDENPCEDIMGKYKFLRYVWFSLALLGWFEVLTILTSYKL 446
594 LPPVAFVAGALGANTLIGSGLASVDALTFQGSSEKARWETGLGRATGLAVLGS 653
447 TVP-RLKCNLAFAFDCMGLYLLLASVDLYTQSEYNNHAIQMGCGCTAGFTVPAS 505
654 EASVLLTLAAVQCSVSVACVRAVGRKSPSLGSVRAGVLCGLAAGLAALPLASVGEYGA 713
506 ELSTVTLITVTLERWATITFAMHIDRKIRLWHAYVIMLGWCCFLALLPLVGISSYAK 565
714 SPLCLPYAPBEGQPALGFTVALVMNNSFCFLVAGAYIKLYCDLPR-----GDFAVWD 768
566 VSLCLPM--DTEPLALAVIILVILNITAFITVACVAKIYITVNPYNGD---KD 619
769 CAMRIVNALIADGLIGPVAFVFLSASMLGLEPVYPEAVKSVLLVLPCLNPLTLV 828
620 TRIARMAVLIITDEMCAPISFYALSALMKNPLITVNSKILLVLEPLNSCANPELYA 679
829 LFNPHRDLRLRPAGDSGLAVYAAAGELKSSCDST 867
680 IFTKAFQRDVFMLSFGICKRQAGAYRQQ-RVSSKNSIT 717
RESULT 15
LSHR_MOUSE
ID LSHR_MOUSE STANDARD; PRT; 700 AA.
AC P30730;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lutropin-Choriogonadotropic hormone receptor precursor (LH/CG-R)
 DE (LSH-R) (luteinizing hormone receptor).
 GN LHCR OR LHR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID:10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92165799; PubMed=1311310;
 RA Gudermann T., Birnbaumer M., Birnbaumer L.;
 RT "Evidence for dual coupling of the murine luteinizing hormone
 RT receptor to adenylyl cyclase and phosphoinositide breakdown and Ca2+
 RT mobilization. Studies with the cloned murine luteinizing hormone
 RT receptor expressed in L cells.";
 RL J. Biol. Chem. 267:4479-4488(1992).
 RN [2]
 RP SEQUENCE OF 1-58 FROM N.A.
 RX MEDLINE=93093308; PubMed=1459341;
 RA Huhtaniemi I.T., Eskola V., Pakarinen P., Matikainen T., Sprengel R.;
 RT "The murine luteinizing hormone and follicle-stimulating hormone
 RT receptor genes: transcription initiation sites, putative promoter
 RT sequences and promoter activity.";
 RL Mol. Cell. Endocrinol. 88:55-66(1992).
 CC -1- FUNCTION: RECEPTOR FOR LUTROPIN-CHORIOGONADOTROPIC HORMONE.
 CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
 CC ACTIVATE ADENYLYL CYCLASE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC FSH/LSH/SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 4 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S49753; AAB24402.1; -;
 DR EMBL: M81310; AAA39432.1; -;
 DR EMBL: M87571; AAA39433.1; -;
 DR PIR: A42395; A42395.
 DR HSSP: P22888; 1LUT.
 DR GCRDB: GCR 0189; -;
 DR GCRDB: GCR 0305; -;
 DR MGD: MGI:96783; lhcgr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 3.
 DR PRINTS: PR00373; GLYCHROMONR.
 DR PRINTS: PR01144; LSHRECEPTOR.
 DR SMART: SMO013; LRRNT; 1.
 DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECPT_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 26
 FT CHAIN 27 700
 FT DOMAIN 27 362
 FT TRANSSEM 363 390
 FT DOMAIN 391 399
 FT TRANSSEM 400 422
 FT DOMAIN 423 443
 FT TRANSSEM 444 466
 FT DOMAIN 467 486
 FT TRANSSEM 487 509
 FT DOMAIN 510 529
 FT TRANSSEM 530 551

FT DOMAIN 552 574 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 575 598 6 (POTENTIAL).
 FT DOMAIN 599 609 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 610 631 7 (POTENTIAL).
 FT DOMAIN 632 700 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 52 75 LRR 1.
 FT REPEAT 126 150 LRR 2.
 FT REPEAT 176 200 LRR 3.
 FT REPEAT 225 248 LRR 4.
 FT DISULFID 443 518 BY SIMILARITY.
 FT CARBOHYD 103 103 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 295 295 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 317 317 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SO SEQUENCE 700 AA; 78214 MW; 8A6840A011E014 CRC64;
 Query Match 9.3%; Score 472; DB 1; Length 700;
 Best Local Similarity 21.3%; Pred. No. 6.2e-25;
 Matches 197; Conservative 147; Mismatches 331; Indels 250; Gaps 31;
 QY 5 PGLRLMILCALCASRRAGAPQPGPACAPCHQEDGIMLSADCELSAVPGDL 64
 DB 6 PALQLLVLAIVLKQSLHSPE--LSGSRCEPCDCAPDGA-----RCGP-- 50
 QY 65 DPLAVLDLNNMNLTELPGLFHHLRFLBELRLSGNHSHINGOAFSGYS-LKIIMLGN 123
 DB 51 -PRAGLRLSLTYLP-----YKVIYSQAFRGINEVVKIISQS 87
 QY 124 NQLGIPAEALMELPELSQSLRLDANLISLVPSEFEGSLSLHMLMDNA-LTEIPVRAL 182
 DB 88 DSLK-----RIEAN-----AFDNLNLSEILIONTKNLIIEPGAF 123
 QY 183 NNPLALQMTALNRISHIIPYAFONLSL-VYLHLNN-RQHGTHSEGLHNETLD 240
 DB 124 TNLPLKLTSLKNGIRLPDVSKISSEFNFLICDNLTYTTPGNAFGQMN-ESTT 182
 QY 241 LNVNKLQEPVAIRTLGRLOELGFHNNIKAIPEAFMGNPLDTTHFDNPLQFVGSA 300
 DB 183 LK-----LYGNFEEVQSAFNCTTIL-SLELKEN----- 211
 QY 301 FOYLPRKLTSLNGAMDIOEPFDLKGTTSLTEILLTRAGIRLLPSGMCQDLRLVLELS 360
 DB 212 -IYLEKMSGTGQATG-----PS-----ILDVS 234
 QY 361 HNOIEPLSLRCKLEBIGLOHNRINWEIGADTFESQSLQALDLSWNAIRSIHPAEST 420
 DB 235 STKLQALPS-----HG-----LESIQTL-----IATSSYS- 259
 QY 421 LHSIVKLDLTDNQLTTLPLAGIGLMHLKLGNALSQAFSKDSPKRLIEVPAVQCC 480
 DB 260 -----LKTLP-----SREKTSLLVALITLYPSHCC 284
 QY 481 PYGMCAFFKASQGEADLHLDDESSKRLGLLAROAEHN-YDQDLDELQ-----E 533
 DB 285 AFRNLPR-----KEQNFSPSIFENFSKQCESTVREANNETLVSAPRENEMLSGWDY 336
 QY 534 MEDSKPHNSVOCSPRPGFKCEYLFESWGRFLAWAIVLVSLONGVLLTVFAGGAP 593
 DB 337 YDFCSP-KTLOCTEPDPAFNPCEDEIMGYAFLLVLLINILAIIFGN-LTVLPLVLLTSRYK 394
 QY 594 LPYKFFVAGALAGANTLTGICSGLLASVDALTFGOFSEYGAWEFGGLGCRATGFLAVLGS 653
 DB 395 LTVREFLMCNLSFADFQGLVILLIASVDSQTKGYYNHAIDMQLGSGCSAAGFTTVRAS 454
 QY 654 EASVILLTLAAVQSVSVYCARVKGSPSLGVSRAVGLCUALAGIAALPLASVGEYGA 713
 DB 455 ELASYITLVITLERMHTITVAVDQKRLRLRAIRIMAGWTFSTLMTPLPLVGVSSYK 514
 QY 714 SPLCLPAAPREGQPAALGFTVALVMANSFCFLVNAVAYITKLYCDLPRQDFEAV-WDCAMV 772
 DB 515 VSICLPM--DVESITSGYIILILLNVAVVICACVRIYFAVQNBELTAPNRDQTKIA 572

```
QY 773 RHVAMLIFFADGLLYCPVAFLESPASMLGIFPVTPPEAVKSVLLVLPPLPACLNPLLYLLENP 832
Db 573 KKMALILFTDFTCMAPISFFAISAAEKVPLITYNSKVLVLYFPVNSCANPFLYAVFTK 632
QY 833 HFRDDLRLRPRAAGDSGLPLAYAAAGLEKSSCDSTQALVAFSDVDLILERSAGRPGL 892
Db 633 AFQBDFFLLLSRFG-----CCKHRAELYRRKEFSACTFNKNGFPRSSK 676
QY 893 TYGPPS-----VLLISCQOPGAPRL 912
Db 677 ----PSQALKLSTIVHCQOPTPRV 697
```

Search completed: October 23, 2002, 10:42:49
Job time : 26 secs

✓

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2002, 10:40:52 ; Search time 42 Seconds
(without alignments)
3983.002 Million cell updates/sec

Title: us-09-851-595-11

Perfect score: 5068
Sequence: 1 MSPSPGRLALMLCALCASR.....GGLSGGCGFQPSGLAFASHV 967

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.yeast:*
16: sp.bacteriophage:*
17: sp.archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4842	95.5	928	4	Q9BYD7
2	4092	80.7	828	4	Q9HBX8
3	2525.5	49.8	907	4	Q75473
4	2514.5	49.6	907	4	Q9UP75
5	2464.5	48.6	907	11	Q9Z1P4
6	2023.5	40.0	951	11	Q9Z2H4
7	2023.5	39.9	951	4	Q9BXB1
8	2012.5	39.7	951	4	Q9BYD1
9	1206	23.8	230	4	Q96K69
10	1095.5	21.6	1360	5	Q9ND11
11	1017.5	20.1	1012	5	Q95Y16
12	1017.5	20.1	1280	5	Q95Y17
13	1010.5	19.9	1300	5	Q9NKD6
14	941	18.6	1050	5	Q9BN18
15	546	10.8	701	13	Q9DGC6
16	514.5	10.2	724	13	Q9PYV0

17	512.5	10.1	688	11	Q64183
18	510.5	10.1	693	13	Q9DGC5
19	503	9.9	779	13	Q918N7
20	491.5	9.7	829	5	Q9VEG4
21	490.5	9.7	831	5	Q94979
22	486.5	9.6	658	13	Q9PVN9
23	486.5	9.6	778	13	Q98TP4
24	486.5	9.6	764	11	Q9DP97
25	484.5	9.6	1091	11	P70193
26	477	9.4	763	6	Q9BGN4
27	473	9.3	1093	4	Q96JAL
28	472.5	9.3	793	13	Q91948
29	472	9.3	814	13	Q91949
30	472	9.3	1094	4	Q9BYB8
31	463	9.1	737	11	Q91Z25
32	462.5	9.1	764	6	Q9BGS6
33	459.5	9.1	603	11	Q70211
34	459	8.9	696	13	Q9DGE5
35	452	8.9	739	6	Q9BGS5
36	447	8.8	662	13	Q9BT84
37	445	8.8	662	13	Q9PWL6
38	442.5	8.7	757	4	Q9BHX9
39	439.5	8.7	603	11	Q9PBL7
40	439.5	8.7	699	4	Q15996
41	438.5	8.7	687	11	Q9J1L0
42	435.5	8.6	696	13	Q9BT85
43	434.5	8.6	701	4	Q14751
44	433.5	8.6	570	11	Q9DBB9
45	432.5	8.5	601	13	Q42500

ALIGNMENTS

RESULT 1

Q9BYD7 PRELIMINARY: PRT: 928 AA.

AC Q9BYD7;
DT 01-JUN-2001 (TREMUREL, 17, Created)
DT 01-JUN-2001 (TREMUREL, 17, Last sequence update)
DT 01-DEC-2001 (TREMUREL, 19, Last annotation update)
DE VTS20631 (FRAGMENT).
GN VTS20631.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki H., Hayashi A., Kozuma S., Saito T.;
RT "a member of g-protein coupled receptor family."
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB049405; BAB39854.1; --
DR HSSP: P23945; 1XUN.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_in.
DR Pfam: PF00560; LRR_15.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR_9.
DR SMART: SM00369; LRR_TYR_14.
FT NON_TER
SQ
SEQUENCE 928 AA: 100487 MW: 403364ADEA89C463 CRC64:

Query Match 95.5%; Score 4842; DB 4; Length 928;
Best local similarity 100.0%; Pred. No. 0;
Matches 928; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 HCGEDGIMLSADCSSELGSAVPGDLPFLAYLDLSNNLTTELPGLEFHRLFEELRLSG 99
DB 1 HCGEDGIMLSADCSSELGSAVPGDLPFLAYLDLSNNLTTELPGLEFHRLFEELRLSG 60

```

OY 100 NMLSHIPGQAFSGLYSLKILMIONNOIGIPAEALMELPSIOSLRFLDANLSLVERSE 159
    |||||
Db 61 NMLSHIPGQAFSGLYSLKILMIONNOIGIPAEALMELPSIOSLRFLDANLSLVERSE 120
OY 160 GISSLRHMLDNLALTEIPVRLNNLPAIQAMTLALNRIISHIPDYAFONTLSVLVLIHN 219
    |||||
Db 121 GISSLRHMLDNLALTEIPVRLNNLPAIQAMTLALNRIISHIPDYAFONTLSVLVLIHN 180
OY 220 NRIQIHGTHSEGLNLETLDTNKNLOEPVATRTGRODELGFHNNNIAIPEKAFMG 279
    |||||
Db 181 NRIQIHGTHSEGLNLETLDTNKNLOEPVATRTGRODELGFHNNNIAIPEKAFMG 240
OY 280 NPLIOTIFHYNPIQFNGRSAPFOYLPKLTLSLNGAMDIOEPDLKGTSTSEILTRAG 339
    |||||
Db 241 NPLIOTIFHYNPIQFNGRSAPFOYLPKLTLSLNGAMDIOEPDLKGTSTSEILTRAG 300
OY 340 IRLPSCMCQOLPRLRVLEISHNOIEELPSLRQKLEELGLOHNRIMEIGADTFSSLS 399
    |||||
Db 301 IRLPSCMCQOLPRLRVLEISHNOIEELPSLRQKLEELGLOHNRIMEIGADTFSSLS 360
OY 400 LQALDLSWMAIRSIHPEAFSTLSVLKDLTDNQLTTPLAGIGGIMHLTKGNLSQA 459
    |||||
Db 361 LQALDLSWMAIRSIHPEAFSTLSVLKDLTDNQLTTPLAGIGGIMHLTKGNLSQA 420
OY 460 FSKDSFKRLTEVPYAYOCCPYGMCASFKASGQWEAEDLHDDERSKRPLGLARQA 519
    |||||
Db 421 FSKDSFKRLTEVPYAYOCCPYGMCASFKASGQWEAEDLHDDERSKRPLGLARQA 480
OY 520 ENHYDDDLDELQLEMEDSKRHPVSQSPYGPCKPCEYLFESWGILAWAYIVLSVLCN 579
    |||||
Db 481 ENHYDDDLDELQLEMEDSKRHPVSQSPYGPCKPCEYLFESWGILAWAYIVLSVLCN 540
OY 580 GIVLTVFAGGAPRLPRPVFVVGALVAGANTLQISGCLLASVDALTFQGFSEGARWETG 639
    |||||
Db 541 GIVLTVFAGGAPRLPRPVFVVGALVAGANTLQISGCLLASVDALTFQGFSEGARWETG 600
OY 640 LGCRAATGFLAVIGSEASVLLTLAAVQCSVSVSVARAYKSSISGVRAGVIGCLALAGL 699
    |||||
Db 601 LGCRAATGFLAVIGSEASVLLTLAAVQCSVSVSVARAYKSSISGVRAGVIGCLALAGL 660
OY 700 AALPLASVGEYGAFLCLPYAPREGOPALGFTVALVMNSFCFLVVAAGATIKLYCDLP 759
    |||||
Db 661 AALPLASVGEYGAFLCLPYAPREGOPALGFTVALVMNSFCFLVVAAGATIKLYCDLP 720
OY 760 RGDFAVMDCAVRRVAVMLIFADGILLYCPVAFSLFASMLGFPYTPPEAVKSVLLVPLP 819
    |||||
Db 721 RGDFAVMDCAVRRVAVMLIFADGILLYCPVAFSLFASMLGFPYTPPEAVKSVLLVPLP 780
OY 820 ACLNPLLLLRNPHRRDLRLRPAGDSGPLAYAAAGELKSSGCDSTQALVAFSDVLI 879
    |||||
Db 781 ACLNPLLLLRNPHRRDLRLRPAGDSGPLAYAAAGELKSSGCDSTQALVAFSDVLI 840
OY 880 LEASBAGRPGLGTYGFPVSYTLISQOQAGAPRLGSHKVEPEGNHFNQPNPSMDGELLRL 939
    |||||
Db 841 LEASBAGRPGLGTYGFPVSYTLISQOQAGAPRLGSHKVEPEGNHFNQPNPSMDGELLRL 900
OY 940 AEGSTPAGGIGSGGGGPOPSGLAFASHY 967
    |||||
Db 901 AEGSTPAGGIGSGGGGPOPSGLAFASHY 928

```

RESULT 2

```

OY 09HBX8 PRELIMINARY; PRT; 828 AA.
AC 09HBX8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-DEC-2001 (Tremblrel. 16, Last sequence update)
DE LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-COUPLED RECEPTOR 6
GN LGR6
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; PubMed=10935549;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA van der Spek P.J., van Duijn M., Huseh A.J.;
RT "The three subfamilies of leucine-rich repeat-containing G protein-
RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT signalling mechanism for LGR7."
RL MOL. Endocrinol. 14:1257-1271(2000).
DR EMBL; AF190501; AAG17168.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR; 10.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEUCRHRPT.
DR SMART; SM00370; LRR; 5.
DR SMART; SM00369; LRR_type; 10.
KW Receptor.
FT NON_TER
SQ
SEQUENCE 828 AA; 89387 MW; 12D9251EA442D8B4 CRC64;
Query Match
Best Local Similarity 80.7%; Score 4092; DB 4; Length 828;
Matches 794; Conservative 0; Mismatches 1; Indels 96; Gaps 1;
OY 72 DISMNNLEIQDGLFHHRLFLBELRLSGNHLSHIPGQAFSGLYSLKILMIONNOIGIPA 131
    |||||
Db 29 DISMNNLEIQDGLFHHRLFLBELRLSGNHLSHIPGQAFSGLYSLKILMIONNOIGIPA 88
OY 132 EALWELPSIOSLRFLDANLSLVERSEGLSSLNHMLDNLALTEIPVRLNNLPAIQAM 191
    |||||
Db 89 EALWELPSIOSLRFLDANLSLVERSEGLSSLNHMLDNLALTEIPVRLNNLPAIQAM 99
OY 192 TLALNRISHIPDYAFONTLSVLVLIHNRIOLGTHSEGLNLETLDTNKNLOEPV 251
    |||||
Db 100 TLALNRISHIPDYAFONTLSVLVLIHNRIOLGTHSEGLNLETLDTNKNLOEPV 112
OY 252 AIRTLGLOELGFHNNNIAIPEKAFMGNPLQTIHYDNPIDFVGRSAFOYLPKLTLS 311
    |||||
Db 113 AIRTLGLOELGFHNNNIAIPEKAFMGNPLQTIHYDNPIDFVGRSAFOYLPKLTLS 172
OY 312 LNGAMDIOEPDLKGTSTSEILTLTRAGIRLPLPSGMCQOLPRLRVLEISHNOIEELPSLI 371
    |||||
Db 173 LNGAMDIOEPDLKGTSTSEILTLTRAGIRLPLPSGMCQOLPRLRVLEISHNOIEELPSLI 232
OY 372 RQCKLEELGLOHNRIMEIGADTFSSLSQALDLSWMAIRSIHPEAFSTLSVLKDLTD 431
    |||||
Db 233 RQCKLEELGLOHNRIMEIGADTFSSLSQALDLSWMAIRSIHPEAFSTLSVLKDLTD 292
OY 432 NQITTLPLAGLGIMHLTKGNLSQAQSFKRLTLEVPYAYOCCPYGMCASFKA 491
    |||||
Db 293 NQITTLPLAGLGIMHLTKGNLSQAQSFKRLTLEVPYAYOCCPYGMCASFKA 352
OY 492 SGQWEAEDLHDDERSKRPLGLLAROENHYDDDLDELQLEMEDSKRHPVSQSPYGP 551
    |||||
Db 353 SGQWEAEDLHDDERSKRPLGLLAROENHYDDDLDELQLEMEDSKRHPVSQSPYGP 412
OY 552 FKPCYELFESWGIRLAWAYIVLSVLCNGVILTYFAGGAPRLPRPVKRVAGALGANTLT 611
    |||||
Db 413 FKPCYELFESWGIRLAWAYIVLSVLCNGVILTYFAGGAPRLPRPVKRVAGALGANTLT 472
OY 612 GISGCLLSVALTFGQSEYGARWETGLGRATGFLAVIGSEASVLLTLAAVQCSVSV 671
    |||||
Db 473 GISGCLLSVALTFGQSEYGARWETGLGRATGFLAVIGSEASVLLTLAAVQCSVSV 532
OY 672 SCVRAAYKSPSLGYSVRAGVLCGLAAGLAAALPLASVGEYGAFLCLPYAPREGOPALG 731
    |||||
Db 533 SCVRAAYKSPSLGYSVRAGVLCGLAAGLAAALPLASVGEYGAFLCLPYAPREGOPALG 592

```


[illegible]

```

RESULT 7
Q9BXBI PRELIMINARY: PRT: 951 AA.
AC Q9BXBI:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE G-PROTEIN COUPLED RECEPTOR.
GN GPR48.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21294803; PubMed-11401528;
RA Loh E.D., Brousseau S.R., Kolakowski L.F.;
RT "Molecular Characterization of a Novel Glycoprotein Hormone G-Protein-
RT Coupled Receptor."
RL Biochem. Biophys. Res. Commun. 282:757-764(2001).
DR EMBL: AF346711; AAK31153.1; -.
DR EMBL: AF346710; AAK31153.1; JOINED.
DR EMBL: AF346710; AAK31153.1; JOINED.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR.
DR InterPro: IPR003592; LRR_Nterm.
DR InterPro: IPR003591; LRR_out.
DR InterPro: IPR000822; znf-C2H2.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR.15.
DR Pfam: PF01462; LRRNT.1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR.6.
DR SMART: SM00013; LRRNT.1.
DR SMART: SM00369; LRR_TYP.15.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
DR KMC
DR RECEPTOR.
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;

Query Match 39.9%; Score 2023.5; DB 4; Length 951;
Best Local Similarity 46.3%; Pred. No. 2,4e-144;
Matches 432; Conservative 130; Mismatches 303; Indels 69; Gaps 12;

QY 1 MSPRGLALMICALCSRRAGAPQEGPPTACAPCHCEQEDGIMLSADCSGLISAY 60
DB 1 MGPGLG-----LC--FLALGLGSGAGPSGAAPPLCAAPCSCDGD--RRVDCSGKGLTAV 51
QY 61 PGDDLPFLAYLDLSNNNTLELPGLFHLRLFEELRSGNHLSHIPGQAFSGLYSLKIM 120
DB 52 PEGLSAFQALDISNNNTLPDEDAFKNPFLLEQLQLAGNDLSFTHPKLSGLKELKYL 111
QY 121 LONNOLGIRPEALWELPSLOSRLDANLISLVPERSEGLSLRHLMLDNDALTEIPR 180
DB 112 LONNOLKIPSEALRGSLASOLSLRDANHLITSVPEDSEGLVQLRHLMLDNDLSLEEVYH 171
QY 181 ALNLPLAOLATLALNRISHIPDAFONLSLVYLHLNNNIOLHGLTHSEGLANLEFLD 240
DB 172 PLNLPLTQALTLANKTISIPDAFNLSSLVVHLHNNNRIRLSQHCFGDLNLEFLD 231
QY 241 LNVKLOEFPVAITLQLOELGFHNNNIKAIPKAFMGNPLDTIHFYNDPIQVNGSA 300
DB 232 LSYNNLGEFFPAIKARSLKELGFHSNISIVIPGADGDNPLTLTHLYDNPISFVNGSA 291
QY 301 FOYLPKHTLSLNGAMIOEPDLKGTTSSEILTLFRAGRILPSPGCOOLPRLVLELS 360
DB 292 FHLNLSLHSLVIRKASWQVPMNLGTVHLSESLTLTGTKISSIPNNLCQKQKMLRTDLS 351
QY 361 HNOIEELPSLRCKOKELEILQHNRIWEIGADFPSSLSQALDLSMNAIRSHPEAFST 420
DB 352 YNNIRLPSNGCHALEISLQKNQITQIKEGTFQGLISRLIRLDSKNLILHEIHSRAFAF 411

```

```

QY 421 LHSIVKLDITDNOITLPLAGIGLMHLKGNLALSQAFSKDSPEKURILEPYAYQCC 480
DB 412 LQPTNLVDVSEFNEILTSFPEGLNGLNQILKLVGNFKLEALAKDFVNRSLSTVPAYQCC 471
QY 481 PYGMCASFEEKASGOWEAEDLHDD-----EESKRLPGLILARAENHYDODDELQLEMD 536
DB 472 AFWGCDST-----ANLTEDNSLQDHSVAQEKTAADANVTSTLENEHSOI----- 518
QY 537 SKPHSVQCSPPPGPFKCEYLFESWGIRLAWYAVLLSVLNGLVLLTFVAGPAPLPP 596
DB 519 -----IHCPTSTGAFKCEYLLGSMIRLFWYFIFVALFENLVLITFASCTG-LPS 572
QY 597 VKFVGAAGANTLTLGICGLASVDALTFQGFSEYKARWETGLACRATGLAVIGSEAS 656
DB 573 SKFLIGLSTVSNLPMGITYTGLITFLDAVSWKRFEEFGTWBETGSCCKVAGFLAVFSSESA 632
QY 657 VLLTLAAVQCSVSVSCVRAVYKSPSLGSRVAGVGLGIALAGLAALPLASVGEASPL 716
DB 633 IFLLMLATVERSLSAKDIMKKNHLLKQFVVALLAFGLATVAGCPPLFIRGEYSASPL 692
QY 717 CLPYAPPEGOPALGFYVYALVMNSFCFLVYVAGYIKYCDLPBGDEAVWDCAMVRHA 776
DB 693 CLPF--PTGETPSLGFVTYVLLNSLAFLLMAVITYTKLYCELEKEDLSSENSOSMIKHA 750
QY 777 WLTFADGLYCPVAFSLFASNLGLFPYTPKAVSVLLVPLPACLPPLLYLTFNPRFD 836
DB 751 WLTFNCIFFCFPAVAFSFAPLITRISPELMSVTILFFPLPCLNPLVTFNPRFKE 810
QY 837 D---LRLRPRGDSPLVYAAAGLEK-----SSCDSTQALVAFSDV 876
DB 811 DWKLKRRVTKKSGSVSVSTSSQGCLEDPFYDCGYSHLQGLNLTVCDCESFLTKPV 870
QY 877 D---LLELSEAGRPPLLETGYFPVTLISCOOP 907
DB 871 SKCHLI-----KSHSCPALAVASQCP 920

RESULT 8
Q9NYD1 PRELIMINARY: PRT: 951 AA.
AC Q9NYD1:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE G-PROTEIN COUPLED RECEPTOR 48.
GN GPR48.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Loh E.D., Brousseau S.R., Kolakowski L.F., Jr.;
RT "Molecular Characterization of a Novel Human Glycoprotein Hormone G-
RT Protein Coupled Receptor."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF257182; AAF68989.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_TYP.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PF00560; LRR.15.
DR Pfam: PF01462; LRRNT.1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR.6.
DR SMART: SM00013; LRRNT.1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.

```


DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 951 AA; 104358 MW; 88DA6C67F6ABA1FE CRC64;

Query Match 39.7%; Score 2012.5; DB 4; Length 951;
 Best Local Similarity 46.0%; Pred. No. 1,6e-143;
 Matches 430; Conservative 130; Mismatches 305; Indels 69; Gaps 12;

QY 1 MESPRLRALMIALCALCASRRAGGAPQPGPGPACAPPCQDQDGLMLSDSCSELGTSAY 60
 DB 1 MESPRLC---LC--FLAGLGLSAGSPSGAPPLCAPCSCDGD---RRVDCSGKGLTAV 51
 QY 61 PGDLDELTAVALDLSMNNLTLPGLFPHHLELRLSGNHLSHIPQAFSGYSLKIM 120
 DB 52 PEGLSAFQALDLSMNNITQLPEDAFKPNFLELQLAGNDLSIFHKALSGELKLVLT 111
 QY 121 LQNNQAGIPAEALMELPISQSLRLDANLISLVPERSFEGISIRHMLDNLTEIPVR 180
 DB 112 LQNNQKTVPESEAIRGLSALQSLRLDANHTSVPEDESFELVQLRHMLDNLSTEYPVH 171
 QY 181 ALNNPALQAMTLALNRISHIIPYAPQNLISVYLHNNRICHGTHSEGLHNETLD 240
 DB 172 PLNLPLQLQTLTALNKKISSIPDAFNLSSLVYLHNNKIRGLSQHCEFDGLDNLLETLD 231
 QY 241 LNNKLOEFPVALRTGLQELGFHNNNNKAIPEKAFMGNPLQTIHFYNDPIQFVGSA 300
 DB 232 LSTNNLGEFQAIKARSLKELGFHSNISVITPDGADGNPLRTIHLNDPLSFVGN 291
 QY 301 FOYLPKRLTSLNGAMDIOEPDLKGTSLLEITLRAGIRLLPSGMCQQLPRLVLELS 360
 DB 292 SHNLSDHSLVIRGASWVOQFPNLGTGVHLESLLTGTKISSIPNNLCQEQKMLRTLDLS 351
 QY 361 HNOIEELPSLRQKKEIEIQHNRIMEIGADMFESQSLQALDLSNNAIRSHIEPAST 420
 DB 352 YNNIRDLPSFNGCHALEISLQRMQYQIKEGTFQGLSKRLDLSRKLHEIHSRAVAT 411
 QY 421 LHSYVKLDLTDNQLTTPPLAGLGIMHLKGNLALSQAESKDSFPRKRLILEVYAAQCC 480
 DB 412 LGPTINDVGFENELTSPTEGPNGLNDKLGNFKLALAKFVNLRSLSVYAAQCC 471
 QY 481 PYGMCASFRRASGQWEADLHLD---ESSKRPGLLARQAENHYDQDLDELQLEMED 536
 DB 472 AFMCDSY---ANLNTEDNSLDHSAOEKGTADAANVTSTLENEHSQI----- 518
 QY 537 SKRPVSOCPTPGPKPCETLFEPMGIRLAVAIVLISVLCNGLVLTVAAGPAPLP 596
 DB 519 -----IHCPTSTGAFFPCETLFGSMMLRLTWFLVLFVLFNLLVLTTFASCTSLPS 572
 QY 597 VKFVGAAGANTLTGISCGLASVDALTFQFSEYGARWETGLGRATGFLAVLGSSEAS 656
 DB 573 SKLPIGLISVSNLPMGIYTGITFLDAVSWGRFAEPFGIMWETSGCKVAGFLAVFSSSA 632
 QY 657 VLLITLAAVQCSVSCVRAVAKSPISGVSAGVGLCLALAGLAALPLASVGEYASPL 716
 DB 633 IFLLMLATVERSLSAKDIMKNGKSNHLKQFRVALLSAFLGATVAGCPPLFRHGEYSAPL 692
 QY 717 CLYAPRPGQPAALGTFTVAIVLMNSFCFLVYAGVYIKYCDLPBGDFEAVADCMVHRVA 776
 DB 693 CLPPE--PTEGPTSLGFTVTLLVLSLAFVLTLYTKYCLEBEDISENSOSSITKIVA 750
 QY 777 WLIFADGLLYCPVAFISFASMLGFVPTPEAVKSVLLVLPCLINPLVLTLPNPHFRD 836
 DB 751 WLITFNCIFPCVPAFFGAPPLITAISSIPKMSVTLIFPLPACINLVLTVPFNPKKE 810
 QY 837 D---LRLRPACDGGFLAVAAAGELK-----SCDSQALVAFSDV 876
 DB 811 DMKLKRRVTKKSGSVSYSSISOGGCEODFYDDCGMYSHLQGNLYVDCCESEFLTKPV 870
 QY 877 D---LILEASAGRPGLGETYGFPSVTLISCOOP 907
 DB 871 SKRHLY-----KSHSCPALAVASCORP 892

RESULT 9

Q96K69

ID Q96K69

PRELIMINARY;

PRT; 230 AA.

AC Q96K69;

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE CDNA FL114471 FIS, CLONE MAMMAL1001030, WEAKLY SIMILAR TO

DE LUTROPIN-CHORIOGONADOTROPIC HORMONE RECEPTOR.

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=MAMMARY GLAND;

RA Tsogal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,

RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,

RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,

RA Makatsuo A., Nakamura Y., Nagahari K., Masubo Y., Oshima A.;

RT "MEDO human CDNA sequencing project.";

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK027377; BAB55071.1; -.

SQ SEQUENCE 230 AA; 24403 MW; 3BF762BA34A9D14F CRC64;

Query Match 23.8%; Score 1206; DB 4; Length 230;

Best Local Similarity 99.6%; Pred. No. 2.3e-83;

Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 738 MANSFCELVVAGAYIKLYCDLPRGDPEAVWDCAMVRHVAWILFADGLLYCPVAELSFASM 797

DB 1 MANSFCELVVAGAYIKLYCDLPRGDPEAVWDCAMVRHVAWILFADGLLYCPVAELSFASM 60

QY 798 LGLFPVPEAVKSVLLVLPAPCLNPLLYLFLNPHFDDLRRLRPAGDSGGLAYAAAG 857

DB 61 LGLFPVPEAVKSVLLVLPAPCLNPLLYLFLNPHFDDLRRLRPAGDSGGLAYAAAG 120

QY 858 ELEKSSCDSTQALVAFSVVDLILEASEAGRPPELETYGFPSVTLISCOOPAPRIEGSHC 917

DB 121 ELEKSSCDSTQALVAFSVVDLILEASEAGRPPELETYGFPSVTLISCOOPAPRIEGSHC 180

QY 918 VEPGNNHGNPQPSMDGELLRAEGSTPAGGGLSGGGGFGPQSLAFASHV 967

DB 181 VEPGNNHGNPQPSMDGELLRAEGSTPAGGGLSGGGGFGPQSLAFASHV 230

RESULT 10

Q9NDI1

ID Q9NDI1

PRELIMINARY;

PRT; 1360 AA.

AC Q9NDI1;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE GLYCOPROTEIN HORMONE RECEPTOR II.

GN RK OR BG:DS00180.13 OR CG8930.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CANTON S.; TISSUE=WHOLE ANIMAL;

RX MEDLINE=20359836; PubMed=10899142;

RA Eriksen K.K., Hauser F., Schott M., Pedersen K.-M., Soendergaard L.,

RA Grimmelikhuijzen C.J.P.;

RT "Molecular Cloning, Genomic Organization, Developmental Regulation,

RT and a Knock-Out Mutant of a Novel Leu-Rich Repeats-Containing G

RL Genome Res. 10:924-938(2000).

DR EMBL; AF142343; AAF6608.1; -.

DR HSSP; Q57815; ID3Y.

DR Flybase; FBgn0003255; rK.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_1p.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 14.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR SMART: SM00370; LRR; 2.
 DR SMART: SM00369; LRR_Typ; 5.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KM Receptor.
 SQ SEQUENCE 1360 AA; 150731 MW; 7D435155B4F6F612 CRC64;

Query Match 21.6%; Score 1095.5; DB 5; Length 1360;
 Best local similarity 30.0%; Pred. No. 7,1e-74;
 Matches 292; Conservative 157; Mismatches 396; Indels 127; Gaps 18;

QY 39 CHCOEDGIMLSADCELSGLSAVPGDLDPLTAYLIDLSMNNITELQPLGFHLRLEELRLS 98
 DB 152 CHCTGSELEVLRLSCRGIGLAVPVNLPNEVVLDGNNNLTKLEANSFPMARLEDLTIS 211
 QY 99 GNHLSHIGQAFSGISYSLKILMLQNNQGGTAEALMEPISQSLRLDANLISLVPERSE 158
 DB 212 DNSTIIMDPNAFYGLAKLRSLQNGKSLPPOSFOGLAOLTSIQNGNALVSLDGDL 271
 QY 159 EGLSSIRHLMDNALTPEIPVALNNPLALQMTLALNRISHIPDYAFQNLGSLVYLHL 218
 DB 272 GHLQKRLRLLEGNFYRIPTNALAGRLLEALNGLSNTLTITINDEDFRMPRLVYLK 331
 QY 219 NNRIOHIGTSHFEGIANLEETLDNLYNKLQEFPAIRTGRLQELGFHNNNIKAIPKARM 278
 DB 332 RNOIMKISAGALKNTALKVLELDNLISSEGLSKISQDELSTSRMLRMINDTELP 391
 QY 279 GNPILQITHEFDNPQIFGRSAFOYLRPKLHTLSLNGAMDIOEPRLDKTSTLEILTLTA 338
 DB 392 RS--WQMDMPRANPLSTISAGFRGMSKRLKLTLDVTLRSFPELEACHALEILKLDRA 449
 QY 339 GIRLPSGMCQOLPRLRVLELSHNOIEELPSLHRC-----Q 374
 DB 450 GIOEPANLCKQTPRLKLELTKNSLRIPNLSGCRDLRLDLSNQIKQKPPNGIK 509
 QY 375 KLEITGLOHNRIMEIGADTFQSLSQALDLSWNAIRSTHPEAFSLHSVLKIDLTNDOL 434
 DB 510 QLANDLLSYNRKIKALPDAFOGIRPKLQDLEGENEISYHKBASFQETALIEDLNGNIF 569
 QY 435 TTLPLAGIGGMHKKINGNLALSQAFSKDSPKRLALEVPYAYQCC---PYGMCASFPA 491
 DB 570 PELPESGLRALHLHKTENNPKLREFPPPTPRIQTLILSYAHCCAFPLVAMSSQKKT 629
 QY 492 S-----GQWEAEDLHIDDESSK 509
 DB 630 SOVGEAVLPSPDAEFDMTLMNNSMNIWPMHNLKSGASMDHWETA-IPNFEQDLOT 688
 QY 510 RPLGLARQENHYDQIDE-----IQLENEDESKPRH SVQCSPTPGPFKPC 555
 DB 689 QTGQOIAITSYEEFEEDHVDGSPATGYGFTGLFSGMSTEDPQ--GSVOCLPMPGFLPC 747
 QY 556 EYLEESGIRLAVVAIVLISVLCNGIVLLTFAGGAPRLPYKFFVVGAMAGANTLIGTIC 615
 DB 748 ADLEFDMWTLRCGVAVVFLSLGNGTVFVLIC--SRSMKDVPRFLVCIINAADFPMGIVL 806
 QY 616 GLLASVADLTFQGFSEYGARWETGLGCRATGFLAVLGSEASVLLTLTAAVQCSVSVCR 675
 DB 807 GIATVDAATLGERBMAIPQMGVILQOLSGFLAVLSELS--YTLAVITLTERNYATH 864
 QY 676 AYKGSPELGSYRAGVLCGL--ALAGLAALPLASVEYGASPLCLPYAPPEGOPALGFT 733
 DB 865 AIIHLNKLRLSKQAGYINSVGWVFALIMALPDLGVSDYRKRAVCPLETTTS--PASTLV 923
 QY 734 VALVMNNSFCFLVAVAGYIKLYCDLPDFEAVV---DCAVVRHVAKLIIFADGLIXCPA 790
 DB 924 ISLAFITGCAFLITLMGCTLAKYMAI-RG--SOAMVNTDSRIAKRMALLVFTDFCLWSPIA 980

QY 791 FLTSASMLGLEPVPYBEAKSVILVLPRLPACINPLILYLFNFHFRDRLRLRPR----- 844
 DB 981 FFSITAFIQLDLSLEQKIFTFVFLPLNCSNPLVIMTKQFKDVCVTLKHEESRV 1040
 QY 845 AGDSGPLAAVAAAGELEKSSCDSTQALVAFSDVDL---ILLEASAGRPPLLETLYGPPSYTL 901
 DB 1041 VGGSGPGRGAVARTKRG-----DLPPPLPAAVAHPPGCR-----CLNM 1081
 QY 902 ISCGQPGAPRL 913
 DB 1082 LPSEMPNWHKME 1093

RESULT 11

ID 095Y16 PRELIMINARY; PRT; 1012 AA.
 AC 095Y16;
 DT 01-DEC-2001 (Tremblrel, 19, Created)
 DT 01-DEC-2001 (Tremblrel, 19, last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, last annotation update)
 DE GLYCOPROTEIN HORMONE RECEPTOR.
 GN APGPHR'.
 OS Asterina pectinifera (Starfish).
 OC Asterozoa; Echinodermata; Eleutherozoa; Asterozoa;
 OC Eukaryota; Metazoa; Valvataceae; Valvatida; Asterinidae; Asterina.
 OC NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita M., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.;
 RT "DNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish Asterina
 RT pectinifera.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061862; BAB8209.1; -.
 KW Receptor.
 SQ SEQUENCE 1012 AA; 112623 MW; 52A70E7A88C4E0A CRC64;

Query Match 20.1%; Score 1017.5; DB 5; Length 1012;
 Best local similarity 30.2%; Pred. No. 3,7e-68;
 Matches 269; Conservative 132; Mismatches 360; Indels 131; Gaps 18;

QY 9 ALWLCALC-----ASRRAGAPQPGPACPARCHQCEGIMLSADCELSAVPGD 63
 DB 12 SYVLLLLLCIRVYISVSCDGTGFLCPTLCC---CH---DG-GHNVCSTRNLTDVPAS 64
 QY 64 LDPILAVLIDSMNNITELQPLGFHLRLEELRLSGNLSHSHIPEQAFGLYSKLIMLQNL 123
 DB 65 LIGITETLISFNNSISILPADAFHRLPRDLTLILIGNLSLTDKNVFGRLNLDLNLK 124
 QY 124 NOLGTPAEALW--ELPSLSIRLDANLISVPERSPGSSSLHIMLMDNALTEIPVA 181
 DB 125 NRFQOYPRKAFRRNDLNLKRLHDSNMWIREVPADAFNLTALHNLNDHQLSEVPTAA 184
 QY 182 LNNLPALQAMTLALNRISHIPDYAFQNLTSVLYLHLNHNRIQHIGTSHFEGIANLETDL 241
 DB 185 LHLHLSNRILHLEHNSIPVYBDHAFANSHLIELILHNKKTTHLSAFAFALPMLMLE- 243
 QY 242 NYKKLOEFPVAIRTLGRLQELGFHNNNIKAIPKAFGNPNLQTIHFYDNPQVGRSAF 301
 DB 244 -----FLGNSITSTAH-----TAF 257
 QY 302 QYPLKHTTSLNGAMDIOEPRLDKTSTLETLTLTRAGIRLRLPSGMCQOLPRLVLELSH 361
 DB 258 RNIPALRNLYLEKKNLSVPPDLTGTSLEHLGIERCSLRATIPANFCDDMGLSTLNHN 317
 QY 362 NQIEELPSLHRCQKLE-----EIGHQHNRIWEIGADTFSQL 397
 DB 318 NLEGGPSSLKSGSKLVLHGTNKLTSLEQPPSGLDIVDQILLENDISITPADAQSL 377
 QY 398 SSLQALDLSNNAIRSHIPEAFSLHSVLKIDLTNDQTLTLPPLAGLGLMHLKGNLALS 457
 DB 378 SHLDLTLSSNNTIREIDSOAFACPTSLQYLDLSNNSPVLPTAGIQMLKIRTYDNLQLE 437

QY 458 QAFKSDSPFKRLILEVPAVYOCPPYGMCAFFKASGQWEADLHDDDESSKRPGLGLAR 517
 DB 438 DEPPSELPTSTETATATAPYHCERYIELAEYLS-----LADRPNISETTYWASG 488
 QY 518 QAENHYDDDELQLEMEDS-----KHPVSQCSPTPGPEKP 554
 DB 489 SVPDYNNMTFEDINSWSIDSIFFGSLSIGSPTYLSGNTSRLVPH-NISCRPKGPEMP 547
 QY 555 CEYLFESMGIRLAWAIVLLSVLCNGVLLVFPAGGPAPLPVKYVGAIAAGANTLTGIS 614
 DB 548 CMLDFGSMPLRIGVWLFLAIGNAIVFVIYVSHTRMDVP-REFLGNLAFADFLGVY 606
 QY 615 CGLLASVDALTFGQSEYKARWETGLCRATGFLAVLGSEAVLLTTLTAAVOCSVSVCV 674
 DB 607 LGFLAGVDTSTLGVFRKRGARQWLSAGCRLAGFLAVFSESESYTLSTLTLEFPAIKHA 666
 QY 675 RAYGKSPSLGSRVAGVIGCLA--LAGLAALPLASVGEYASPLCLPYAPPEGOPALGF 732
 DB 667 LHLEKRMKL--PHAITVMCFGMIFSVTAALPLVNVASHYHRVPCLPF--DVDITVAKY 722
 QY 733 TVALVMNSFCFLVAVAGYIKLYCDLPRGDFEAWDC--AMVRVAVMLIFADGLLYCPV 789
 DB 723 VGSILILNLAVIIMACYASITYAI-OGSH--AMNCNDSRVARRMSLIVFTDFACWAPI 779
 QY 790 AFLSPASMLGFPVTPPEAVKSVLLVPLPACLNPLLYLLENPHRDDLRL 841
 DB 780 AFFSLTAARGLRLISLDGAKVLTIFVLPLNSCANPFLYTLTKOKKCKTI 831

RESULT 12

ID Q95Y17 PRELIMINARY; PRT: 1280 AA.

AC Q95Y17
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLYCOPROTEIN HORMONE RECEPTOR.
 GN APGHR.
 OS Asterina pectinifera (Starfish).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
 CC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
 OX NCBI_TaxID=7594;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mita N., Hirai T., Oba Y., Yoshikuni M., Nagahama Y.
 RT "cDNA cloning and functional analysis of a novel member of the
 RT glycoprotein hormone receptor family from a starfish Asterina
 RT pectinifera.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB061861; BAB68208.1;
 KW Receptor.
 SO SEQUENCE 1280 AA; 141700 MW; 0AB0ECC0DD880BA CRC64;

Query Match 20.1%; Score 1017.5; DB 5; Length 1280;
 Best Local Similarity 30.2%; Pred. No. 5.2e-68;
 Matches 269; Conservative 132; Mismatches 360; Indels 131; Gaps 18;

QY 9 AMLWCAALC-----ASRRAGAPQGPQPTACAPACHCOEDGIMLSADCSGLASVAPD 63
 DB 12 SYVLLLLCLRVVSSRCVGDGFLCPGLCC--CH--DG-GTHVNCSTRRLNDVPS 64
 QY 64 LDPLTAVYLDLSNNLTLEQPLFHLRFEELRLSGNHLSTHPCGAFSLYSKTIIMLN 123
 DB 65 LIGITETIDLSNNITISILPADAFRHLRDLTILILGNRLSTDKNVFRGLRINDTLNKL 124
 QY 124 NOLGSIPEALM--ELPSQSLRLDANLISLVPESEFGLSIRLMLDNDNALTEIPVA 181
 DB 125 NRPQVPRKAFRNDLANLRKLTLDNSNIREVPADAFNMLTALHNLNDHNOISEVPTAA 184
 QY 182 LNNPLALQAMTALNRISHIIPYAFONLTSVYVHLNRRIGHGHSPEGANLETLDD 241
 DB 185 LHLTSLRLTLHLSHSIPVYPDPAFAENSHLIELILRHKKITHLSAHAPAGLPNMLLE- 243

QY 242 NNNKLEPPVALIRTLGRLOELGFHNNNIKAJPEKAMGNPLQTLHFYDNPLOVGRSAF 301
 DB 244 -----FLNGSTISAH-----TAF 257
 QY 302 QYLPRLHTLSLNGAMDIOEPDLKGTSTLEITLTRAGIRLLPSGCOOLPLRLVLESH 361
 DB 258 RNLPLRLNVLILEVNNLSVFPDLTGTSTLEHIGRCSLRAPANFCONMTGLTSLNHN 317
 QY 362 NOIEELPSLHROQKLE-----EIGLQHNRIWEIGADTFESQL 397
 DB 318 NLIEGLPSLSCSSLSKVLHLGNTKLTLSLEGOPSGIHLDYDLOLENDISTYIPALAFOSL 377
 QY 398 SSLQALDLSWNAIRSIHPEAFSTLSLVKLDLTNQLTTLPLAGLGLMLKLKGNLALS 457
 DB 378 SHLDLTLSLNNITIRELDGAFAPCTSLQYLDLSNNSFPVLPFAGLOMLKRTYDNEQE 437
 QY 458 QAFKSDSPFKRLILEVPAVYOCPPYGMCAFFKASGQWEADLHDDDESSKRPGLGLAR 517
 DB 438 DEPPSELPTSTETATATAPYHCERYIELAEYLS-----LADRPNISETTYWASG 488
 QY 518 QAENHYDDDELQLEMEDS-----KHPVSQCSPTPGPEKP 554
 DB 489 SVPDYNNMTFEDINSWSIDSIFFGSLSIGSPTYLSGNTSRLVPH-NISCRPKGPEMP 547
 QY 555 CEYLFESMGIRLAWAIVLLSVLCNGVLLVFPAGGPAPLPVKYVGAIAAGANTLTGIS 614
 DB 548 CMLDFGSMPLRIGVWLFLAIGNAIVFVIYVSHTRMDVP-REFLGNLAFADFLGVY 606
 QY 615 CGLLASVDALTFGQSEYKARWETGLCRATGFLAVLGSEAVLLTTLTAAVOCSVSVCV 674
 DB 607 LGFLAGVDTSTLGVFRKRGARQWLSAGCRLAGFLAVFSESESYTLSTLTLEFPAIKHA 666
 QY 675 RAYGKSPSLGSRVAGVIGCLA--LAGLAALPLASVGEYASPLCLPYAPPEGOPALGF 732
 DB 667 LHLEKRMKL--PHAITVMCFGMIFSVTAALPLVNVASHYHRVPCLPF--DVDITVAKY 722
 QY 733 TVALVMNSFCFLVAVAGYIKLYCDLPRGDFEAWDC--AMVRVAVMLIFADGLLYCPV 789
 DB 723 VGSILILNLAVIIMACYASITYAI-OGSH--AMNCNDSRVARRMSLIVFTDFACWAPI 779
 QY 790 AFLSPASMLGFPVTPPEAVKSVLLVPLPACLNPLLYLLENPHRDDLRL 841
 DB 780 AFFSLTAARGLRLISLDGAKVLTIFVLPLNSCANPFLYTLTKOKKCKTI 831

RESULT 13

ID Q9NRD6 PRELIMINARY; PRT: 1300 AA.

AC Q9NRD6; Q9YJ03;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 144.0 KDA PROTEIN (RK GENE PRODUCT).
 DE RK OR BG:DS00180.13 OR CG8930.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Y, AND CN BW SP;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazek R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 RA Celinker S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-kb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [2]

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishit S., Hsu S.Y., Zell K., Hsueh A.J.;
 RT "Characterization of two fly LGR (leucine-rich repeat-containing G
 RT protein-coupled receptor) proteins homologous to vertebrate
 RT glycoprotein hormone receptors: constitutively activation of wild type
 RT fly LGR but not LGR2 in transfected mammalian cells."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF274591; MAK0808.1;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR_10.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR SMART: SM00370; LRR; 8.
 DR SMART: SM00369; LRR_typ; 10.
 DR PROSITE: PS50262; G_PROTEIN_RECIP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 1050 AA; 117707 MW; 35D71260A8B4BF99 CRC64;
 Query Match 18.6%; Score 941; DB 5; Length 1050;
 Best Local Similarity 27.8%; Pred. No. 2,4e-62;
 Matches 259; Conservative 148; Mismatches 349; Indels 176; Gaps 21;
 OY 104 HIRGO--AFSGVLSLKL-----MLQNNQGGIPAEALWELP 138
 DB 2 HLPNNYLFSEHIFRSLFSPVLYGFCNGYKKRKTNNPHQNLGNALVSDGDLGHQ 61
 OY 139 SIQSLRLDANLISVPERSEFGLSLRLMDNALTEIPYRALNMLPALQAMTLALNRI 198
 DB 62 KLRLRLRLEGNL-----FYRIPTNALAGRLTEALNIGSNIL 97
 OY 199 SHIDYAFQNTLSVYLHLNHNRIOLGTHSEFGLHNTELDLNNKIQEPPVALRTIGR 258
 DB 98 TINDDEPRPNMLVLILKLRQIKRISAGALKNTALKVELDNLSSLEPESKLSIQ 157
 OY 259 LOELGFHNHNKAIPKAFMGNPILQTIHFNDPIQFVGRSAFOYLKHLTSLNGAMD 318
 DB 158 LOELSTNRLRLMIDTELPS--WQMDMKNPLSTISPGAFRMSKRLKILSDVRL 215
 OY 319 QEPDLKGTSLLEITLRAGIRLLPSGMCQQLPRLVLELSHNOIEELPSLHRC----- 373
 DB 216 RSFPELECHALEILKLRAGIQEVPANLCROTPLKSLKLTNSIKRIPNLSSCRDLRL 275
 OY 374 -----QKLEELGLOHNRIMEIGADTFSSLSIQALDLSMNAIRSIH 414
 DB 276 LDLSNQIEKIQGKPPSGIKQNLDLILSTNRKALKAPDQFQGIPIKQLDLGNEISYTH 335
 OY 415 PRAFSSTLSVYKLDITDQTLTPLAGLGMLKLGMLKGNALSAQSFPRKRLILEVP 474
 DB 336 KRAFGSGTALDELNLGNINFPPLPESGLRALLHLTKTFNNPKLRREPPTPTPRIGLILS 395
 OY 475 YAYQCC---PYGMCASFKA----- 492
 DB 396 YAYHCAFLPLVYAMSSOKKTSQVOEAVLPSPDAEPTTLNMNMMNIMWPMHNLKQOLGA 455
 OY 493 ---GQMEADLILDDSESKRPLGLLAROENHYDQDLDE-----LQLEME 535
 DB 456 SMHDPRETA- INFNEQLOTOGQIATISYMEEFEEHYSGPATYGGCTGLFSMSME 514
 OY 536 DSKPSPVOCSPTPGPFKPCXYLFESWGIRLAVMATLVLLSVLCNGLVLLTVFAPGAPRL 595
 DB 515 DQGP- GSVQCLPMPGPFELCADFLDMWTLRGVWVYVFLSLILNGTVYVFLC- SRSKMD 572
 OY 596 PVKFWVGAITAGANTLTGICGLASDALTFQGFSEYSGARWETGLCGRATGFLAVLGSEA 655
 DB 573 VRFELVCLNLAADDFEWGIYLGILAIYDAATLGEFRMAFIPWMSVLCOLSGFLAVLSSEL 632

OY 656 SVLLTLAAVQCSVSCVRATGKSPSLGSYRAGVLCGL--ALAGLAALPLASVGEYA 713
 DB 633 SV-YTLAIVTLERNAVITRAIHLNKRSLKQAGIYMSVGVFRLIAMAPLIVGSYRK 690
 OY 714 SPICLPYAPREGOPALGFVALYVMNSFCFLVVAAGVYIKYCDLPQDGEAVW---DCA 770
 DB 691 FAVOLPPEYTTG-PASLTLYVISTLMFINGCAFLTLMGCLKMYAI-RG--SQAMTNSR 746
 OY 771 MVRHVALIPADGILYCPVAFSLFASMLGLPPYRPAVKSVLYVLRPLACPLNLYLLE 830
 DB 747 IAKRMALVDFDFLCWSPFIAFSSITAIPLGLISLEQAKITVVLPLNSCNPFYAIIM 806
 OY 831 NPHFRDLRLRPR-----AGDSPLAAGAELKSSCDSTOALVAFSDVDL---ILE 881
 DB 807 TKQFKKDCVTLKAFEEBSRVVGGGPGRGAVARTKRG-----DLPPILP 852
 OY 882 ASEAGRPGLTYGFPYSYTLISCOQPGARLE 913
 DB 853 AAAYAHPPGCR-----CLRMPLSEMPNMHME 879
 RESULT 15
 Q9DGC6 PRELIMINARY; PRT; 701 AA.
 AC Q9DGC6;
 DT 01-MAR-2001 (TREMBLrel. 16, created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE GONADOTROPIN RECEPTOR 1.
 GN TGNH-RL.
 OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
 CC Clariidae; Oreochromis.
 OX NCBI_TaxID=8128;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oba Y., Hirai T., Yoshitani Y., Yao Z., Nagahama Y.;
 RT "Subtilin gonadotropin receptor 1";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB041762; BAB16106.1; --
 DR HSSP: P23945; 1XON.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR InterPro: IPR001611; LRR.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 5.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECIP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECIP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 701 AA; 78453 MW; B3D78465CA56410A CRC64;
 Query Match 10.8%; Score 546; DB 13; Length 701;
 Best Local Similarity 26.8%; Pred. No. 1.1e-32;
 Matches 177; Conservative 112; Mismatches 310; Indels 62; Gaps 21;
 OY 226 GTHSF-EGHNLNTELDLNNKLOEPPV-AIRTLGRQELGFHNHN- IKAIPKAFMGNP 282
 DB 51 GYTAFFSNISNAQCLEVKQTHQIHEIQGGTSSLOHLMELTISNDLESIGAFSGLP 110
 OY 283 LQTIHFYDN-PIQFVGRSAFOYLKHLTSLNGAMDIOEP-PP-LKGTSLLEITLRAG 339
 DB 111 LTKILISKNAALNIGAFVFSNLPSELSEIITKSKHLSEFIHPDAFRMARLRFLLTSNTG 170
 OY 340 IRLPBGMCQQLPRLVLELSH-QIEELPS--LHRC-QKLEELGLOHNRIMEIGADTF 394
 DB 171 LRIFPFSKIHSTACPLDLQDNSHIKRVANAFRGICLYCTFPAIRLTRNGIKEVASDAF 230
 OY 395 SOLSSLOALDLSWN-AIRSIHPEAFSTLSVYKLDITDQTLTPLAGLGMLKLGKGN 453
 DB 231 NG-TKMRFLPLGNRQRLTHSPNAFVSSSELVVDVSEFATLSLPDSITLGLKRLIESA 289

